

# Enzymological Analysis of STAT6 Signaling Related Proteins and Its Application to the Drug Discovery and Development

著者（英）	Mariko Hirozane
year	2019
その他のタイトル	STAT6シグナル関連タンパク質の酵素学的解析と新規薬剤開発への応用
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2018
報告番号	12102甲第9053号
URL	<a href="http://doi.org/10.15068/00156352">http://doi.org/10.15068/00156352</a>

Enzymological Analysis of STAT6 Signaling Related Proteins and  
Its Application to the Drug Discovery and Development

January 2019

Mariko Hirozane

Enzymological Analysis of STAT6 Signaling Related Proteins and  
Its Application to the Drug Discovery and Development

A Dissertation Submitted to  
the Graduate School of Life and Environmental Sciences,  
the University of Tsukuba  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy in Biological Science  
(Doctoral Program in Biological Sciences)

Mariko Hirozane

## Table of Contents

Abstract.....	4
Abbreviations.....	6
General Introduction.....	9

### Chapter 1 Identification of PARP14 Inhibitors

#### Using Novel Methods for Detecting Auto-ribosylation

Abstract.....	24
Introduction.....	25
Materials and Methods.....	28
Results.....	34
Discussion.....	39

### Chapter 2 High-throughput Screening to

#### Identify Inhibitors of DEAD Box Helicase DDX41

Abstract.....	52
Introduction.....	53
Materials and Methods.....	55
Results.....	62
Discussion.....	66
General Discussion.....	82
Acknowledgements.....	87
References.....	89

## Abstract

In this research, I focused on the signal transducer and activator of transcription 6 (STAT6) signal in two different pathways of various immune cell types—the interferon-4 (IL-4)/interferon-13 (IL-13) activating pathway in Th2 cells and the DNA virus activating pathway in dendritic cells. IL-4 and STAT6 mediate Th2 responses in T cells and immunoglobulin class switching to IgE in B cells, thus regulating the balance between inflammatory and allergic immune responses. On the other hand, the DNA virus activating pathway in dendritic cells initiates the transcription of genes encoding chemokines such as chemokine (C-C motif) ligand 2 (CCL2) and chemokine (C-C motif) ligand 20 (CCL20). These chemokines especially enhance tumor specific immunity, presumably in a T-lymphocyte-dependent manner, and STAT6 signaling can play an important role in tumor immunity. Poly (ADP-ribose) polymerase 14 (PARP14) and DEAD-Box helicase 41 (DDX41) are promising molecular targets of drug discovery related to STAT6. In this study, I biochemically characterized these two enzymes and identified their selective inhibitors by high-throughput screening (HTS). The enzymes related to STAT6, PARP14, and DDX41, were evaluated by enzymological analysis. First, novel assay methods were established for evaluating the autoribosylation activity of PARP14 using HTMS and immunoradiometric assay. Through these assays, HTS was conducted using a library of more than 500,000 compounds and hit compounds were identified. The co-crystallization of PARP14 and some of the hit compounds revealed that the inhibitors interacted with the nicotinamide adenine dinucleotide (NAD<sup>+</sup>)-binding site of PARP14. Finally, I confirmed that the identified hit compounds were able to interact with intracellular PARP14 using a cell-based protein stabilization assay. Thus, primary candidate compounds were

successfully identified. Secondly, I clarified that the ATPase activities of human DDX41 and R525H are dependent on ATP and dsDNA, either through poly(dG–dC) or poly(dA–dT). Novel inhibitors of DDX41 and R525H were identified by HTS. Hit compounds with selective inhibitory activities were identified using several assays. These results provide an ideal platform for the discovery of PARP14 and DDX41 inhibitors, present new chemical tools for use in investigating the multiple functions of these enzymes, and open a new horizon for drug discovery research related to STAT6 signaling.

## Abbreviations

## Abbreviations

AML	acute myeloid leukemia
BCR	B cell receptor .
CCL	chemokine (C-C motif) ligand
DDX41	DEAD-Box Helicase 41
DTT	dithiothreitol
E. coli	Escherichia coli
HIPK2	delta-1Homeodomain Interacting Protein Kinase 2
HTMS	high throughput mass spectrometry
HTS	high-throughput screening
IC <sub>50</sub>	50% inhibitory concentrations
IFNs	interferons
IRF3	interferon regulatory factor 3
JAK	janus kinase
MDS	myelodysplastic syndrome
NAD <sup>+</sup>	nicotinamide adenine dinucleotide
NF-κB	nuclear factor-kappa B
PARP14	poly (ADP-ribose) polymerase 14
PARP14 <sup>-/-</sup>	poly (ADP-ribose) polymerase 14 knockout
PARPs	poly(ADP-ribose) polymerases
PCR	polymerase chain reaction
PLCD1	phospholipase C, delta-1
R525H	R525H mutant
RORγ	retineic acid receptor related orphan nuclear γ



Runx3 runt related transcription factor 3

SD standard deviation

SDS-PAGE sodium dodecyl sulfate-polyacrylamide gel electrophoresis

STAT signal transducer and activator of transcription

STING stimulator of interferon genes

TBK1 TANK-binding kinase 1

TCR T cell receptor

Treg regulatory T cells

WT wild type

dsDNA double-stranded DNA

## General Introduction

## General Introduction

### Immune system and immune responses

Organisms often encounter pathogens like bacteria or viruses that cause potentially harmful disease. However, they can function properly and live without constantly being sick in most cases because organisms like humans have a multilayered immune system to keep it running without difficulties. Immune responses are mainly categorized into innate immunity and adaptive immunity as shown in Fig. 1 and Table 1 [1, 2].

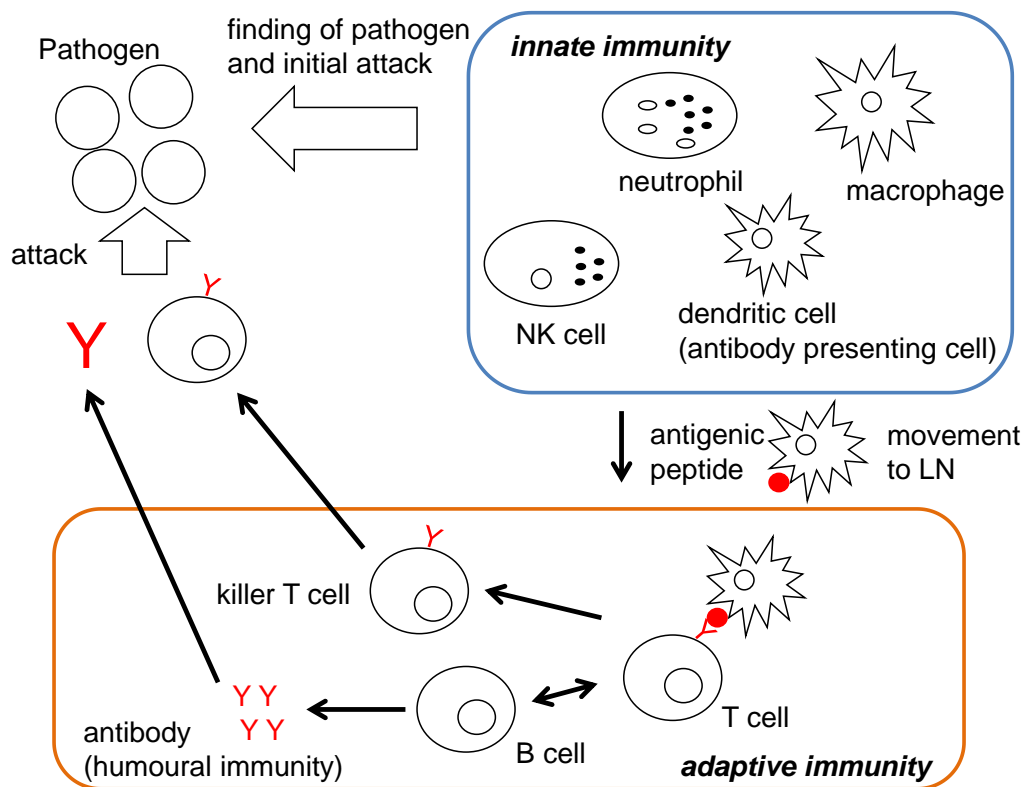


Figure 1. Interactions of innate and adaptive immunity.

Innate immune responses are mediated by leukocytes, macrophages, and dendritic cells, which are collectively called phagocytes because they engulf and kill microbes. In the case of dendritic cells, they have an additional role of presenting antigenic peptides

derived from microbes to T cells [3]. Dendritic cells engulf pathogens, digest them into small peptides, and express the peptide antigens on their cell surface. The dendritic cells then migrate from the infected tissue to the regional lymph nodes, where they present the antigens to naive T cells with parallel receptors. Cells in the innate immune system recognize and respond to pathogens discriminating by self from foreign pathogens. Although the innate immune system is an immediate defense against infection, it is a temporary system until adaptive immune responses can be triggered. In higher metazoans such as vertebrates, adaptive immune responses occur in addition to innate immune responses.

Table 1 Innate and adaptive immunity

	Innate immune system	Adaptive immune system
Property	Fixed in genome	Encoded in gene segments
Receptors	Rearrangement is not necessary	Rearrangement necessary
Distribution	Non-clonal	Clonal
	All cells of a class identical	All cells of a class distinct
Recognition	Conserved molecular patterns (LPS, LTA, mannans, glycans)	Details of molecular structure (proteins, peptides, carbohydrates)
Self-Non-self discrimination	Perfect: selected over evolutionary time	Imperfect: selected in individual somatic cells
Action time	Immediate activation of effectors Co-stimulatory molecules	Delayed activation of effectors Clonal expansion or anergy
Response	Cytokines (IL-1 $\beta$ , IL-6) Chemokines (IL-8)	IL-2 Effector cytokines: (IL-4, IFN- $\gamma$ )

Adaptive immunity is activated by exposure to pathogens, and uses immunological memory to learn about the threat and enhance the immune response unlike the innate immune system. The adaptive immune response is much slower to respond to threats and infections compared to the innate immune response as shown in Table 1, and functions through B cells and T cells. Both cells are lymphocytes derived from specific types of multipotent hematopoietic stem cells in the bone marrow. After they are made in the bone marrow, they are matured and activated. Each cell type follows a different path to attain the mature form.

In the lymphatic system, maturation of B cells begins when naïve B cells encounter an antigen [4-6]. Naïve B cells express one of millions of distinctive antibodies on their cell surface, which can also be called membrane-bound antibodies. When a naïve B cell encounters an antigen that can react with its membrane-bound antibody, it quickly divides in order to become either a memory B cell or plasma B cell. Memory B cells produce antibodies only upon re-stimulation by the specific antigen and their response in conferring protection against reinfection has remained controversial. Plasma B cells secrete antibodies to protect from infection. The secreted antibodies identify free pathogens circulating throughout the body. B cells also express a specialized receptor called the B cell receptor (BCR). BCRs assist with antigen binding, as well as internalization and processing of the antigen. After the antigen is internalized and processed, the B cell can initiate signaling pathways such as cytokine release and communicate with other cells of the immune system.

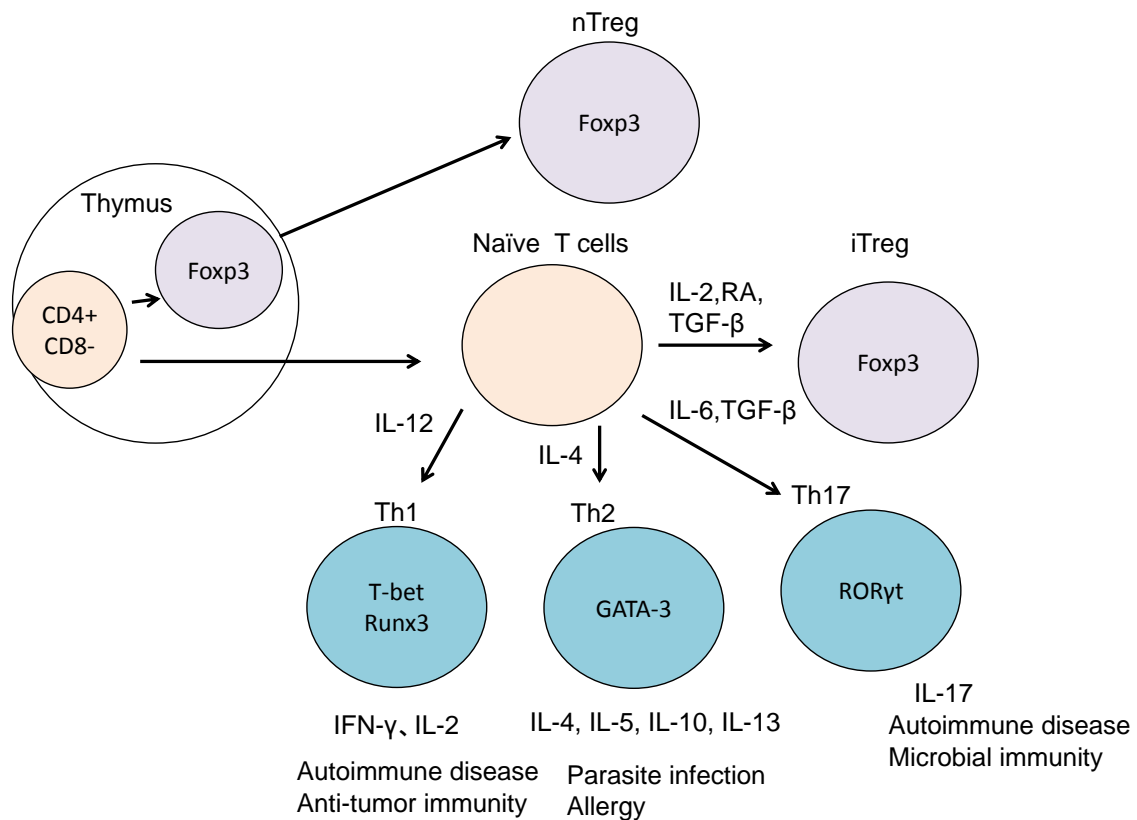


Figure 2 Differentiation of naïve CD4+ T cells into Tregs or effector T cells

T cells also have important roles in the adaptive immune system. Once formed in the bone marrow, T progenitor cells migrate to the thymus for maturation and become T cells. The developing T cells express T cell receptors (TCRs), and either CD4 or CD8 receptors in the thymus [7, 8]. TCRs recognize antigens that are bound to certain receptor molecules and undergo a process called rearrangement, resulting in nearly limitless recombination of a gene that encodes antibodies. Naïve CD4+ T cells can be differentiated into regulatory T cells (Tregs) or effector T cells shown in Fig. 2. The transcription factors T-bet and runt related transcription factor 3 (Runx3), GATA3, or retinoic acid receptor related orphan nuclear  $\gamma$  (ROR- $\gamma$ ) are required for the

differentiation of naive T cells into Th1, Th2, or Th17 cells, respectively. Mature T cells should recognize only foreign antigens in order to mount an appropriate immune response and secrete specific cytokines (Table 2).

Table 2 Examples of cytokines and primary activities

Cytokines	Principal Source	Primary Activity
GM-CSF	Th cells	Growth and differentiation of monocytes and dendritic cells
IL-1 $\alpha$	Macrophages and other antigen presenting cells (APCs)	Costimulation of APCs and T cells, inflammation and fever, acute phase response, hematopoiesis
IL-2	Activated Th1 cells, NK cells	Proliferation of B cells and activated T cells, NK functions
IL-3	Activated T cells	Growth of hematopoietic progenitor cells
IL-4	Activated T cells	B cell proliferation, eosinophil and mast cell growth and function, IgE and class II MHC, expression on B cells, inhibition of monokine production
IL-5	Th2 and mast cells	Eosinophil growth and function
IL-6	Activated Th2 cells, APCs, other somatic cells	Acute phase response, B cell proliferation, thrombopoiesis, synergistic with IL-1 and TNF on T cells
IL-7	Thymic and marrow stromal cells	T and B lymphopoiesis
IL-8	macrophages, somatic cells	Chemoattractant for neutrophils and T cells
IL-9	T cells	Hematopoietic and thymopoietic effects
IL-10	Activated Th2 cells, CD8+ T and B cells, macrophages	Inhibits cytokine production, promotes B cell proliferation and antibody production, suppresses cellular immunity, mast cell growth
IL-11	Stromal cells	Synergistic hematopoietic and thrombopoietic effects
IL-12	B cells, macrophages	Proliferation of NK cells, IFN production, promotes cell-mediated immune functions
IL-13	Th2 cells	IL-4-like activities
IL-18	Macrophages	potent inducer of interferon- $\gamma$ by T cells and NK cells
IFN- $\alpha$	Macrophages, neutrophils and some somatic cells	Antiviral effects, induction of class I MHC on all somatic cells, activation of NK cells and macrophages
IFN- $\beta$		
IFN- $\gamma$	Activated Th1 and NK cells	Induces of class I MHC on all somatic cells, induces class II MHC on APCs and somatic cells, activates macrophages, neutrophils, NK cells, promotes cell-mediated immunity, antiviral effects
MIP-1 $\alpha$	Macrophages	Chemotaxis
MIP-1 $\beta$	Lymphocytes	Chemotaxis
TGF- $\beta$	T cells, monocytes	Chemotaxis, IL-1 synthesis, IgA synthesis, inhibit proliferation
TNF- $\alpha$	macrophages, mast cells, NK cells, sensory neurons	Cell death, inflammation, pain
TNF- $\beta$	Th1 and Tc cells	phagocytosis, NO production, cell death



## STAT pathways activated by cytokines

Cytokines are small secreted proteins released by cells. They have a specific effect on the interactions and communications between immune cells [9, 10]. Cytokine is a general name; other names include lymphokine, monokine, chemokine, and interleukin (IL). Cytokines may react with the cells that secrete them, with nearby cells, or in some instances, with distant cells. They are often produced in a cascade, as one cytokine stimulates its target cells to express additional cytokines. Cytokines exert their functions by binding to specific cellular receptors; therefore both cytokine and receptor expression can determine the nature of an individual cell response, and in a broader context, it can impact on the intensity and duration of the inflammatory response, ultimately affecting the clinical outcome of periodontal/periapical disease [9].

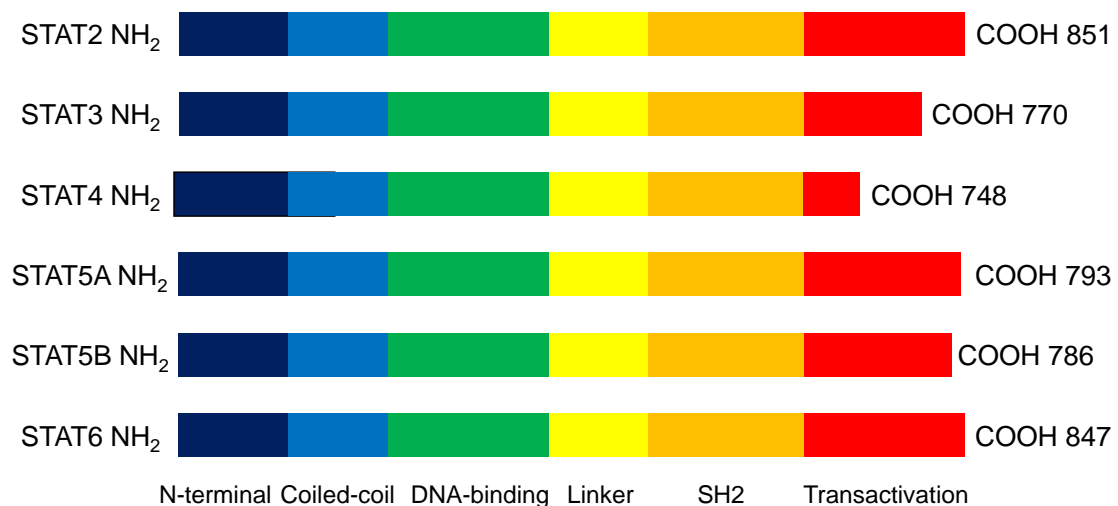


Figure 3 A schematic representation of STAT proteins

Some cytokines, including ILs, interferons (IFNs), and hematopoietic growth factors, activate Janus kinase (JAK)/signal transducers and activators of transcription (STAT) pathway [11-13]. STAT proteins comprise seven members: STAT1, STAT2, STAT3,

STAT4, STAT5A, STAT5B, and STAT6 (Fig. 3). These share five domains structurally, which include an amino-terminal domain, a coiled-coil domain, a DNA-binding domain, an SH2 domain, and a carboxy-terminal transactivation domain. Activation of STAT is promoted as shown in Fig. 4, when ligands such as cytokines bind their receptors expressed on the cellular membrane. The cytokine-receptor interaction induces receptor phosphorylation, which recruits STAT proteins to the phospho-motifs of the receptor. Phosphorylation of the critical tyrosine residue in STAT proteins is then initiated by tyrosine kinases, such as growth factor receptors, JAKs (JAK1, JAK2, JAK3, and Tyk2), and SRC family kinases [14]. Phosphorylation of a tyrosine residue present in the trans activation domain and two phosphorylated STAT monomers dimerize through reciprocal pTyr-SH2 domain interactions, and these STAT-STAT dimers translocate to the nucleus where they bind to specific STAT-response elements in the target gene promoters and regulate transcription. STAT-dependent induction of genes is essential for many physiological functions. Activation of normal STAT signals is controlled by physiologically negative modulators such as suppressors of cytokine signals and protein tyrosine phosphatases, in accordance with normal cellular functions. Although STAT proteins are differentially activated and promote varying cellular processes depending on the ligand and cell types, their normal induction collectively regulates cell growth, differentiation, survival, and apoptosis. Such inductions lead to regulation of inflammatory and immune responses, embryonic development, and mammary gland development. On the other hand, defective or abnormal STAT signaling is associated with various human diseases including susceptibility to infection, immune disorders, many types of cancer, asthma, and allergic diseases. Therefore, STAT signaling is one of the important signals to control human diseases.

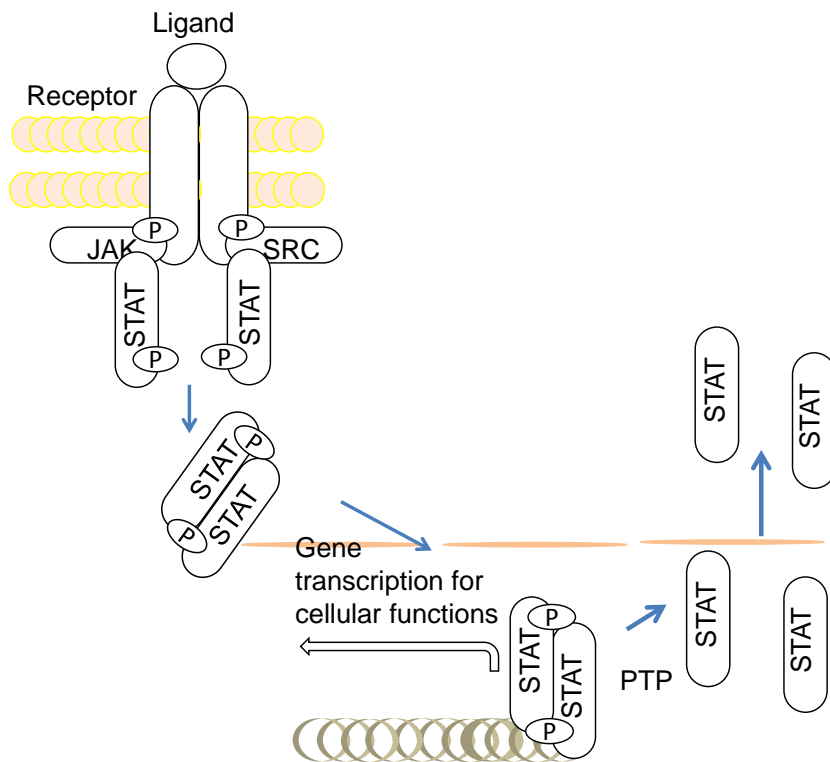


Figure 4 STAT signal pathway

Some studies have indicated that STAT6 signaling plays a predominant role in the immune system. STAT6 signaling was shown to be present in two different pathways in various cell types (Fig. 5): the IL-4/IL-13 activating pathway and the DNA virus activating pathway [15]. IL-4 induces the phosphorylation of IL-4 receptor, which in turn recruits cytosolic STAT6 by its SH2 domain; the recruited STAT6 is phosphorylated at tyrosine 641 (Y641) by JAK1, which results in its dimerization and nuclear translocation to activate the target genes [16-19]. Some of the STAT6-bound regions coincide with various epigenetic marks, and the corresponding genes include IL-4, GATA3, IL-24, phospholipase C, delta-1(PLCD1), and homeodomain interacting

protein kinase 2 (HIPK2). This leads to Th2 polarization. IL-4 and STAT6 mediate Th2 responses in T cells and immunoglobulin class switching to IgE in B cells, thus regulating the balance between inflammatory and allergic immune responses [20-23]. Therefore IL-4 and STAT6 signaling is an attractive therapeutic target pathway. To date, various studies on discovering the modulators of STAT6 signaling have been reported. Some peptides, small molecules, or natural products are reported to inhibit STAT6 dimerization, whereas some small molecules are reported to inhibit JAK1 that phosphorylates STAT6. However, these inhibitors show weak activities against STAT6 protein or nonspecific activities against JAK1 even if the compounds possess high activity. Therefore, discovery of promising novel molecular targets in IL-4 and STAT6 signaling has been desired. Previous studies indicate that poly (ADP-ribose) polymerase 14 (PARP14) enhances IL-4-induced gene expression by interacting with the cytokine induced signal transducer as an activator of STAT6 in B and T cells. In the absence of IL-4, PARP14 binds to STAT6 responsive promoters, and functions as a transcriptional repressor by recruiting HDAC2, HDAC3, and p100 co-factor. On the other hand, PARP14 is replaced in the presence of IL-4 and transcription of STAT6 responsive genes is initiated. Therefore, PARP14 is required for STAT6-dependent gene expression. Moreover, some biological experiments have shown that inhibition of PARP14 possibly prohibits STAT6-related diseases [24, 25]. These studies indicate that PARP14 is an enzyme modulating the IL4 and STAT6 signaling and that PARP14 can be a more suitable drug target molecule compared to STAT6 or JAK, because some inhibitors of PARP1, which belongs to the PARP family similar to PARP14, have been reported and showed selective and high inhibitory activities against PARP1[26-28].

On the other hand, DNA virus infection triggers STAT6 activation in some immune

cells such as dendritic cells [29, 30]. Stimulator of interferon genes (STING), which plays a central role in integrating signals from DNA virus infection, is activated by DNA virus infection and then recruits STAT6. TANK-binding kinase 1 (TBK1) is also recruited by STING to phosphorylate STAT6 on S407, which in turn activates another unidentified tyrosine kinase to phosphorylate STAT6 on Y641, leading to the homodimerization and nucleus translocation of STAT6. The STAT6 dimer then binds to its target sites to initiate the transcription of genes encoding chemokines such as CCL2 and CCL20. These chemokines especially enhance tumor specific immunity, presumably in a T-lymphocyte dependent manner indicating that this STAT6 signal can play an important role in tumor immunity [31, 32]. Research efforts to discover molecules to modulate STING are still ongoing in the field of STAT6 signaling. On the other hand, obtaining small molecules with activities against STING is considered difficult because the number of reported small molecules against STING is extremely small despite its popularity. The reported agonists and antagonists are mainly compounds like cyclic dinucleotides. Recently, DEAD-Box Helicase 41 (DDX41) was discovered as a molecule that modulates STAT6 signaling in accordance with STING [33, 34]. Moreover, a comprehensive genetic analysis identified a somatic mutation in DDX41, which was shown to be a cause of the familial acute myeloid leukemia (AML)/myelodysplastic syndrome (MDS) pedigree.

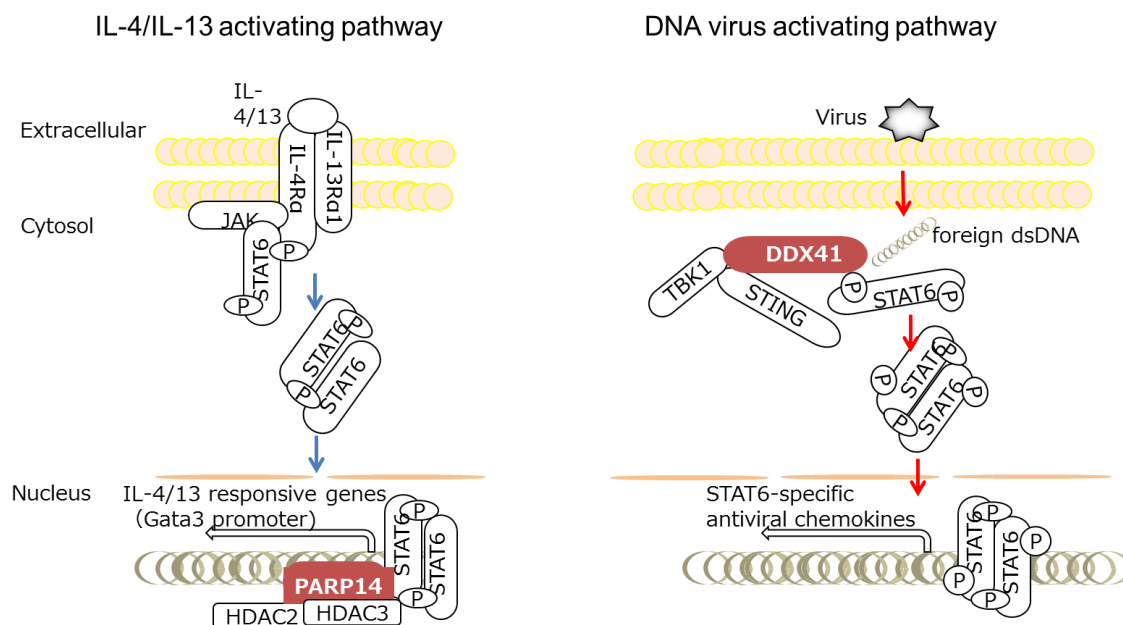


Figure 5 Two different pathways in STAT6 signaling

In this thesis, I have described my studies focusing on PARP14 and DDX41 as significant molecules in STAT6 signaling. It was revealed that these enzymes might play an important role in STAT6- related diseases. However, very little is known about the biochemical properties of these enzymes even though some biological experiments have been reported. Thus, I biochemically characterized these enzymes and identified selective inhibitors against these enzymes by high-throughput screening (HTS).

In the first chapter, I describe the analyses for PARP14. I developed novel assay methods for detecting PARP14 activity. Moreover, I identified potent, selective, and cell-permeable compounds by HTS. The inhibition mechanism of the hit compounds was analyzed by solving the X-ray crystal structure of PARP14 in complex with certain hit compounds. In the second chapter, I describe the analyses for DDX41. I evaluated the ATPase activities of human DDX41 which is dependent on ATP and dsDNA. Then, novel selective inhibitors of DDX41 were identified through HTS.

The results of these analyses provide an ideal platform for the discovery of PARP14 and DDX41 inhibitors, present new chemical tools to investigate the multiple functions of these enzymes, and open a new horizon for drug discovery research related to STAT6 signaling.

Chapter 1 : Identification of PARP14 Inhibitors Using Novel Methods  
for Detecting Auto-ribosylation



## Chapter 1 : Identification of PARP14 Inhibitors Using Novel Methods for Detecting Auto-ribosylation

### Abstract

Poly(ADP-ribose) polymerases (PARPs) use nicotinamide adenine dinucleotide (NAD<sup>+</sup>) as a co-substrate to transfer ADP-ribose when it releases nicotinamide as the metabolized product. Enzymes of the PARP family play key roles in detecting and repairing DNA, modifying chromatin, regulating transcription, controlling energy metabolism, and inducing cell death. PARP14, the original member of the PARP family, has been reported to be associated with the development of inflammatory diseases and various cancer types, making it a potential therapeutic target. In this study, I purified the macrodomain-containing PARP14 enzyme and established an assay for detecting the auto-ribosylation activity of PARP14 using RapidFire high-throughput mass spectrometry and immunoradiometric assay using [<sup>3</sup>H]NAD<sup>+</sup>. Subsequently, I performed high-throughput screening using the assays and identified small-molecule hit compounds, which showed NAD<sup>+</sup>-competitive and PARP14-selective inhibitory activities. Co-crystal structures of PARP14 with certain hit compounds revealed that the inhibitors bind to the NAD<sup>+</sup>-binding site. Finally, I confirmed that the hit compounds interacted with intracellular PARP14 by a cell-based protein stabilization assay. Thus, I successfully identified primary candidate compounds for further investigation.

## Introduction

Poly(ADP-ribose) polymerases (PARPs) are a family that share the PARP motif to catalyze the transfer of ADP-ribose to target proteins [35]. PARPs catalyze the formation of linear or branched poly(ADP-ribose) chains, which are dynamic structures recognized by reader domains and that are broken down by poly(ADP-ribose) glycohydrolases [36]. Seventeen genes in the human genome possess PARP domains and each PARPs are endowed a variety of other functional domains [37]. A subset of the PARP family catalyzes mono-ADP-ribosylation but not chain elongation [38], and functional analyses on such mono-ADP-ribosyltransferases have been performed [39, 40]. Most of these combine an ADP-ribosyltransferase domain with ADP-ribose-binding macrodomains, WWE domains, CCCH-type zinc finger domains, and other protein-protein interaction modules. PARP14 has three macrodomains which is predicted as interacting with ADP-ribose and considered as a mono-ADP-ribosyltransferase [41].

Previous reports indicated that PARP14 enhances IL-4-induced gene expression by interacting with the cytokine induced signal transducer and activator of STAT6 in B and T cells [42, 43]. In the absence of IL-4, PARP14 was found to be bound to STAT6 responsive promoters, and functioned as a transcriptional repressor by recruiting HDAC 2, 3 and p100 co-factor. In the presence of IL-4 the catalytic activity of PARP14 modified the HDACs and the repressive complex was displaced from the promoter to activate transcription. PARP14 is required for STAT6-dependent gene expression. These studies indicated that PARP14 has the ability to bind DNA but the exact sequence to which PARP14 binds is not known. PARP14 protein was predominantly expressed in lymphoid tissues (spleen, thymus, and lymph nodes) [43]. This lymphoid pattern of

expression suggested that PARP14 may function predominantly in the immune system.

The expression profile of genes between wild type (WT) and PARP14 knockout (PARP14<sup>-/-</sup>) Th2 cells were analyzed to define the mechanism PARP14 participates in Th2 cell differentiation and impacts allergic airway disease [42]. The Th2 cytokine genes *Il4*, *Il5* and *Il13* were reduced in PARP14<sup>-/-</sup> Th2 cells. Moreover, the RNA pol II binding profile of genes in Th2 differentiation was analyzed and *Gata3* was reduced due to PARP-14 deficiency. These results suggest that PARP14 may regulate the expression of GATA3 by controlling the binding ability of STAT6 to the *Gata3* promoter [42]. Therefore, PARP14 is considered as a promising drug target for STAT6 related disease. However, the methods for evaluating the activities of PARP14 inhibitors have focused on biophysical assays, such as thermal shift assay, small-molecule microarrays, isothermal titration calorimetry, and differential scanning fluorimetry [44-48]. However, only the enzyme-linked immunosorbent assay has been reported to be suitable for evaluating enzymatic activity, and this method is time consuming because of the requirement for washing steps. Therefore, I here report a novel assay method using RapidFire high throughput mass spectrometry (HTMS) to evaluate PARP14 activity. HTMS is an attractive method for label-free screening of enzyme targets by directly quantifying native, unmodified substrates and products based on their mass-to-charge ratio; it has also recently been successfully used for screening some enzymes[49, 50].

In my study described in Chapter 1, novel assay methods for evaluating the autoribosylation activity of PARP14 by HTMS and immunoradiometric assay were established. Through these assays, I conducted a screening analysis from a library of more than approximately 500,000 compounds and identified hit compounds. Moreover,

I showed that co-crystallization of PARP14 and some hit compounds occurred, indicating that the inhibitors interacted with the nicotinamide adenine dinucleotide (NAD<sup>+</sup>) -binding site. Finally, I confirmed that the identified hit compounds could interact with intracellular PARP14 by a cell-based protein stabilization assay. The primary candidate compounds were successfully identified, so I conclude that these results represent a significant advance in PARP14-related research.

## Materials & Methods

### Phylogenetic analysis of PARP14

By Basic Local Alignment Search Tool Protein (blastp) searches, we collected 38 PARP14 sequences including a human PARP14 from non-redundant protein sequences (nr) database of National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/>) by using the human PARP14 sequence (NP\_060024.2) as a query. Hit sequences that possess three macro domains characteristic PARP14 with E-value less than  $1 \times 10^{-10}$  were selected. Sequences were aligned using Muscle program [51] in SeaView package version 4.6.1 [52]. We selected 1,251 positions of unambiguously aligned by manual and used for phylogenetic analyses. The data matrix for phylogeny were subjected to the RAxML program version 8.2 [53]. GTR +  $\Gamma$ 4 model for amino acid substitution process was applied to infer the maximum likelihood (ML) tree. In the bootstrap analysis, heuristic tree search was performed with a rapid bootstrap algorithm option (-f) for 100 bootstrap replicates. Bootstrap proportion (BP) values were indicated on the corresponding internal branches of the ML tree drawn by the use of FigTree program Version 1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

### Protein Expression and Purification

To express human PARP14 (residues 791–1801) in insect cells, an entry plasmid, pENTRGST(V)/hPARP14 (791–1801), was constructed using a human heart cDNA library (Takara Bio, Kusatsu, Japan) as a template. A DNA fragment encoding human PARP14 (791–1801) was amplified by polymerase chain reaction (PCR) using PrimeSTAR® GXL DNA Polymerase (Takara Bio), and the PCR product was ligated to the EcoR I/Not I site of the pENTR-D-TOPO vector (Thermo Fisher Scientific,

Waltham, MA), preceded by insertion of a sequence for the N-terminal GST-tag. Baculovirus was prepared using the BaculoDirect™ Baculovirus Expression System (Thermo Fisher Scientific) using pENTRGST(V)/hPARP14 (791–1801). Sf9 cells were infected with baculovirus and cultured for 72 h. Recombinant proteins were purified by affinity chromatography and gel filtration chromatography. All purification steps were performed at 4°C. Harvested Sf9 cells were suspended in lysis buffer [50 mM HEPES (pH 8.0), 150 mM NaCl, 3 mM DTT, 20% glycerol, 0.04% Tween 20, 2 mM KCl, and 5 U/mL Benzonase], sonicated for 3 min, and then centrifuged at  $13,000 \times g$  for 40 min at 4°C. The supernatant was incubated with 45 mL of glutathione-Sepharose™4B resin (GE Healthcare, Little Chalfont, UK) equilibrated with buffer A [50 mM HEPES (pH 8.0), 150 mM NaCl, 3 mM DTT, 20% glycerol, 0.04% Tween 20, and 2 mM KCl] for 60 min. The resin was washed with 270 mL of buffer A and then GST-tagged PARP14 (residues 791–1801) was eluted with 20 mM reduced glutathione in the same buffer. The eluate containing GST-tagged PARP14 (791–1801) was concentrated using an Amicon Ultra-15 (Millipore, Watford, UK). GST-tagged PARP14 (791–1801) was further purified by gel filtration chromatography with a HiLoad 26/600 Superdex 200 pg column (GE Healthcare) equilibrated with buffer A.

#### RapidFire Mass Spectrometry

The PARP14 enzyme reaction was conducted with 100 nM GST-tagged PARP14 and 0.5 mM NAD<sup>+</sup> diluted with assay buffer [50 mM HEPES (pH 8.0), 10 mM MgCl<sub>2</sub>, 1 mM DTT, 0.01% Tween20, and 0.01% BSA] in a 384-well polypropylene plate (Corning, Corning, NY) at a final volume of 2 µL. The reaction was performed for 120 min and terminated by adding 60 µL of 4% formic acid containing 6.3 µM

nicotinamide-d4 (C/D/N Isotopes, Quebec, Canada).

HTMS was performed using a RapidFire300 system (Agilent, Santa Clara, CA) coupled to an Agilent6490 triple quadrupole mass spectrometer (Agilent). The samples were loaded onto a Graphitic Carbon cartridge (Agilent) using 0.1% trifluoroacetic acid in deionized water at a flow rate of 1.5 mL/min and eluted to the mass spectrometer using 0.05% trifluoroacetic acid in acetonitrile/acetone/deionized water (70/20/10, v/v/v) at a flow rate of 1.0 mL/min. The sipper was washed to minimize carry over with deionized water followed by acetonitrile. Aspiration time, load/wash time, elution time, and re-equilibration time were adjusted to 350, 3000, 5000, and 1000 ms, respectively, for a cycle time of approximately 10 s. The mass spectrometer equipped with an electro-spray ion source was operated in positive multiple reaction monitoring (MRM) mode. The MRM transitions (Q1/Q3) for nicotinamide as a reaction product and nicotinamide-d4 as an internal standard were set as 123.0/80.0 and 127.0/84.1, respectively. For the NAD<sup>+</sup> competition assay, the samples were measured by an ultra-high-performance liquid chromatography Nexera MP system (Shimadzu, Kyoto, Japan) coupled to Agilent6490.

The inhibitory activity was calculated using the peak area ratio, which is the reaction product divided by its internal standard. I defined the peak area ratio of the reaction without enzyme as 100% inhibitory activity and that of the complete reaction mixture as 0% inhibitory activity. Curve fittings and calculations of IC<sub>50</sub> values were performed with GraphPad Prism V.5 (GraphPad Software, San Diego, CA).

### Immunoradiometric Assay

Assay methods for detecting the auto-ribosylation activity of PARP14 and PARP1

were established using [ $^3\text{H}$ ]NAD $^+$ . The enzyme reaction was performed on a plate coated with anti-GST antibody (GE Healthcare). The PARP14 enzyme reaction was conducted with 30 nM GST-tagged PARP14 and 80 nM [ $^3\text{H}$ ]NAD $^+$  diluted with assay buffer [50 mM HEPES (pH 8.0), 1 mM MgCl $_2$ , 1 mM DTT, 0.01% BSA (Sigma, Darmstadt, Germany)] for 210 min at room temperature. The reaction was terminated by washing with PBST buffer and MicroScint-PS (PerkinElmer, Waltham, MA) was added to the plate. In the case of the PARP1 selectivity assay, the enzyme reaction was conducted with 7.1 nM GST-tagged PARP1 (BPS Bioscience, San Diego, CA), 30 nM [ $^3\text{H}$ ]NAD $^+$ , and 2  $\mu\text{g/mL}$  activated DNA (Trevigen, Gaithersburg, MD) diluted with assay buffer for 210 min at room temperature. Radioactivity formed by autoribosylated PARP1 and PARP14 was read using a TopCount (PerkinElmer). The inhibitory activity was calculated in the same way as for HTMS.

#### Crystallization and Structure Determination of Human PARP14

A clone was made as per a previously published structure reported by the Structural Genomics Consortiu [44]. The region of PARP14 covering residues 1611–1801 was cloned into *Escherichia coli* (*E. coli*) vector containing a kanamycin resistance marker and a tobacco etch virus (TEV) protease cleavage site. *E. coli* Rosetta DE3 strain was scaled up for large-scale expression and purification. Cells were lysed by sonication followed by centrifugation at 17,968 x g for 1 h. The supernatant was loaded on an affinity NiNTA column (Qiagen, Hilden, Germany) and purified using an increasing linear gradient of nickel sulfate (20–200 mM). The eluate was checked for purity and treated with TEV protease followed by size exclusion chromatography on a HiLoad 26/600 Superdex 200 pg (GE Healthcare). After confirmation of the pure fractions,



protein was concentrated to 38.2 mg/ml in buffer [25 mM Tris (pH 7.9), 200 mM NaCl, and 0.5 mM TCEP] and used for structural studies. Initial screening in a PEG Ion screen (Hampton Research, Aliso Viejo, CA) yielded several hits both with apo protein and in the presence of 1 mM concentration of compounds. However, the best diffracting crystals appeared when optimized under the following conditions [20% PEG3350 and 0.2 M sodium-malonate (pH 7.2–7.8)]. Crystals were harvested after cryoprotection with either 15% ethylene glycol or 15% glycerol for 15 min.

Data were collected at Beamline 5.0.3 of the Advanced Light Source at Berkeley (CA, USA). The structure was solved by molecular replacement using Molrep [54] of the CCP4 package [55]. PDBId 3GOY was used as the starting model. This was followed by iterative rounds of manual model building in Coot [56] and restrained refinement using Refmac5 [57]. The ligand atoms and water molecules were fit manually after interpreting the difference map and accounting for the protein atoms.

#### Intracellular PARP14 Stability Assay

The A549 cell line stably expressing NanoLuc-fused PARP14 (PARP14-NanoLuc, residues 1611–1801) was established and cultured in DMEM medium (Wako Pure Chemical Industries, Ltd., Osaka, Japan) supplemented with 10% fetal bovine serum (Corning) and 0.5 mg/mL G-418 (Wako).

After seeding 5,000 cells per 20  $\mu$ L of medium into a white opaque 384-well plate (Corning) overnight, 5  $\mu$ L of test compounds were added to the cell plate, which was then incubated for 6 h at 37°C in 5% CO<sub>2</sub> and 95% humidity. In the final step, 25  $\mu$ L of Nano-Glo Luciferase substrate (Promega, Fitchburg, WI) was added to the assay plate and luminescence was read using an EnVision plate reader (PerkinElmer). The

inhibitory activity was calculated using the luminescence data and curve fittings and calculations of EC<sub>50</sub> values were performed with GraphPad Prism V.5 (GraphPad Software).

## Results

### Phylogenetic Analysis of PARP14

To evaluate what kind of organisms possess PARP14 containing three macro domains, molecular evolutionary analysis was performed. The blastp search using human PARP14 as a query revealed that PARP14 homologs with three macro domains are present only in restricted metazoan lineages, Porifera (sponges), Cnidaria, and Chordata. Although I could not find the PARP14 homologs in other metazoan phyla including Arthropoda, the presence of PARP14 in the earliest branching metazoan phylum, Porifera, suggests the presence of PARP14 in the common ancestor of metazoan lineages. Outside Metazoa I could neither find the PARP14 homologs in choanoflagellates nor in fungi or plants, also suggesting the emergence of PARP14 in the ancestral metazoan lineage. In the ML tree of PARP14 (Figure 6), several lineage specific gene duplications were observed, while in mammals only a single gene was present. On the branch leading to each of the duplicated genes of *Ciona intestinalis*, the evolutionary rate might have been accelerated. Although the putative homologs of STAT6 protein was conserved not only in diverse metazoans but also in amoeba (*Dictyostelium*) [58], the function of the STAT6 homologs except for model animals have not been clarified in detail. However, in the origin of animals the STAT6 signal might have needed to be more complex in order to evolve multicellularity of animals. Considering such situation, PARP14 might have been established in the common ancestor of metazoans as an additional enzyme for regulating the complex STAT6 signal. Since then it has been conserved during the evolution leading to vertebrates, but has been lost in several other metazoan lineages.

### Assay Development for Evaluating PARP14 Auto-ribosylation Activity

To evaluate the auto-ribosylation activities of PARP14, HTMS and immunoradiometric assays were established. The HTMS assay was initially developed for the primary screen. PARP14 was titrated and the reaction was monitored for 120 min (Fig. 7A). The production of nicotinamide by PARP14 increased over time in a linear manner and the  $K_m$  value of  $NAD^+$  was determined to be 1.2 mM (Fig. 7B). The condition for the screen was set to 500  $\mu M$   $NAD^+$ , 0.1  $\mu M$  PARP14, and a reaction time of 120 min. To validate the assay, PJ-34 (Wako), a known PARP family inhibitor, was titrated under the optimized assay conditions and determined to have an  $IC_{50}$  value of 9.0  $\mu M$  (Fig. 7C). High reproducibility of the compound activities was confirmed by HTMS assay (Fig. 7D). Moreover, an immunoradiometric assay was also developed using [ $^3H$ ] $NAD^+$  to detect the PARP14 auto-ribosylation activity. Titration of [ $^3H$ ] $NAD^+$  and PARP14 was monitored for 240 min (Fig. 8A, B), resulting in the assay conditions of 30 nM PARP14, 80 nM [ $^3H$ ] $NAD^+$ , and a reaction time of 210 min to keep a stable signal-to-noise ratio of approximately 3.5. The  $IC_{50}$  of PJ-34 was 2.5  $\mu M$ , which is comparable to the activity detected by HTMS (Fig. 8C).

### Identification and Characterization of Novel Inhibitors

A diverse library of more than approximately 500,000 small-molecule compounds in Takeda Pharmaceutical Company was screened at 3  $\mu M$ , as shown in the work flow (Fig. 9A, B). The initial hit compounds were eliminated and possessed almost the same activities in the HTMS and immunoradiometric assays. To elucidate the mechanism of inhibition, a substrate competition assay was performed. The rates of PARP14 inhibition by the compounds were examined in the two conditions described in the Materials &

Methods. The representative hit compounds listed in Table 3 showed lower IC<sub>50</sub> by a 20-fold increase in NAD<sup>+</sup> concentration (Table 3). The results indicated that these hit compounds possessed competitive inhibitory activity against NAD<sup>+</sup>. These results were supported by the co-crystallization of PARP14 and hit compound, as described below. Moreover, a selectivity assay against PARP1 was conducted. Compound 1 (6-{3-[4-(diphenylmethoxy)piperidin-1-yl]propoxy}[1,2,4]triazolo[4,3-b]pyridazine) showed no activity against PARP1 at a concentration of 30 μM, while compound 2 (2-[(pyridin-4-ylmethyl)amino]-5,6,7,8-tetrahydro[1]benzothieno[2,3-d]pyrimidin-4(3H)-one) has an IC<sub>50</sub> value of 26 μM (Table 3).

#### X-ray Crystal Structure of PARP14 in Complex with Hit Compounds

To elucidate the mode of binding of the inhibitor with PARP14, I obtained X-ray crystal structures of PARP14 in complex with the hit compounds. I selected compound 3 (N-{4-[4-(diphenylmethoxy)piperidin-1-yl]butyl}[1,2,4]triazolo[4,3-b]pyridazin-6-amine) and compound 4 (2-[(1-methylpiperidin-4-yl)methyl]amino}-5,6,7,8-tetrahydro[1]benzothieno[2,3-d]pyrimidin-4(3H)-one), which belong to the same chemotypes as compound 1 and compound 2, respectively. The IC<sub>50</sub> values of compound 3 and compound 4 are 610 nM and 76 nM in the immunoradiometric assay of PARP14. Data collection and refinement statistics of the structures are shown in Table 4. Compound 3 binds to the catalytic site, making several key interactions, mainly with the terminal bicyclic core (Fig. 10A). Two of the triazole nitrogens form hydrogen bonds, one with the Ser1722 side chain hydroxyl and the other with the Gly1683 backbone NH. The bicyclic plane of the

inhibitor stacks parallel to the side chain of Tyr1727. The rest of the aliphatic tail and the terminal phenyl rings are exposed to the solvent. Unsurprisingly, the density in the difference map is lower for this region of the inhibitor. The loop of residues 1699–1710 is disordered in this structure. The two terminal phenyl rings, however, are involved in crystal contact-mediated interactions with an adjacent molecule, where they face residues Asn1720 and Tyr1721.

In the case of compound 4, the planar tricyclic ring binds to the active site sandwiched between the aromatic rings of Tyr1727 and Try1714 (Fig. 10B). The sulfur atom faces the solvent side, while the carbonyl and the pyrimidine nitrogen atom form hydrogen bonds with the backbone NH and carbonyl of Gly1683, respectively. Additionally, the inhibitor carbonyl forms a hydrogen bond with the hydroxyl of Ser1722. The terminal methyl piperidine faces the solvent orienting toward the side chains of Asp1685 and Ser1688. Atom coordinates and structure factors for complexes will be deposited in the Protein Data Bank.

The obtained crystal structures were compared with the crystal structure of PARP1 (PDB ID: 4L6S). An overlay of these structures is shown in Fig. 10C for compound 3 and Fig. 10D for compound 4. It was revealed that the two terminal phenyl rings in compound 3 and the terminal methyl piperidine ring of compound 4 sterically clashed with part of the PARP1 protein. These observations are well consistent with the assay results that hit compounds exhibited good selectivity against PARP1.

#### Detecting Interaction Activity of Intracellular PARP14

To evaluate the cellular activities of the representative hit compounds including compounds 1–4, a novel detection system for quantifying the interaction of transfected

PARP14-NanoLuc with hit compounds inside the cell was established. The accumulation of PARP14-NanoLuc for 6 h after treatment with the compound was detected. A correlation plot of the pIC50 values of enzyme immunoassay compared with the cell-based protein stabilization assay shows a high positive correlation (Fig. 11,  $r=0.82$ ). The hit compounds are cell-permeable and interact with the intracellular PARP14.

## Discussion

In this section, I established assay methods for detecting PARP14 activity. Moreover, I identified potent, selective, and cell-permeable compounds through high-throughput screening (HTS). A strategy for creating potential lead compounds can be developed by solving the X-ray crystal structure of PARP14 when in complex with certain hit compounds.

A variety of assays have been used to identify potential PARP14 inhibitors [45] [46] [47] [48], but the methods of evaluating PARP14 enzymatic activities have been less well studied. I established an HTS assay for PARP14 with HTMS, which is beneficial as it involves a label-free method that can quantify native, unmodified substrates and products based on their mass-to-charge ratio. Macrodomain-containing purified PARP14 was used because it possesses the conserved ADP-ribose binding motif [39] [40] [41], but the differences in activity between non-ribosylated and ribosylated PARP14 were not determined in the auto-ribosylation assay because the enzyme itself can also serve as an acceptor. In the future, it might be necessary to analyze the molecular function of PARP14 by identifying endogenous substrates other than the enzyme itself, before studying the reaction mechanism.

Some compounds were identified through HTS, I also developed a PARP1 selectivity assay to verify the specificity of enzymatic inhibition, and a cell-based protein stabilization assay to estimate cellular activity. The hit compounds possessed potent inhibitory activities, as shown in Table 3, with both competitive activities against  $\text{NAD}^+$  and target interaction activities in the cells. These HTS results indicate that PARP14 is a highly druggable target and hit compounds directly target PARP14. By co-crystallization analysis and  $\text{NAD}^+$  competition assay, I also demonstrated that there



was a prominent binding site for small-molecule inhibitors that overlapped with the NAD<sup>+</sup> binding site. Based on the crystal structural information, it would be possible to guide these hit compounds into good lead compounds as PARP14 inhibitors. Regarding compound 3, the aliphatic tail and the terminal phenyl rings do not show specific interactions with PARP14 and it is suggested that a shorter tail or a smaller terminal group would be sufficient for steric repulsion with PARP1. Taking these findings together, modification of the aliphatic tail and terminal phenyl rings to a smaller substructure would maintain both inhibitory activity and good selectivity against PARP1. In the case of compound 4, it was expected that the introduction of substituents to the terminal piperidine ring that were intended to interact with Asp1685 and Ser1688 would enhance inhibitory activity without losing good selectivity against PARP1.

In another part of the research, I established a method of cell-based monitoring of the interaction between small-molecule inhibitors and intracellular targets using a NanoLuc luciferase detection system (Promega) because the direct marker of PARP14 was unknown. Ligand-induced protein stabilization is a proven phenomenon [59-61], and recent reports have described assay systems using enzyme fragment complementation, time-resolved fluorescence resonance energy transfer, and cellular thermal shift [62-64]. It was confirmed that the compounds with functional cellular or tissue activity could stabilize the intracellular target molecule. Therefore, the cell-based protein stabilization assay of PARP14 can be considered an efficient way to validate target engagement, with NanoLuc luciferase detection being a well-established and robust method that is frequently used in drug screening for high-sensitivity detection [65]. Because NanoLuc luciferase has a low molecular weight (19 kDa), it is well suited for use as a protein fusion tag, allowing luminescence to be associated with the physiological dynamics of

specific intracellular proteins. The protein stabilization induced by small-molecule compounds was measured by quantification of the intracellular transfected PARP14-NanoLuc fusion protein. Some hit compounds identified through HTS showed a correlation for the IC<sub>50</sub> values obtained by enzyme assay and the PARP14 cell-based stabilization assays (Fig. 11).

Finally, it has been reported that PARP14 is expressed in lymphoid organs and lymphocyte cells, and that it has a significant role in the development of allergic inflammation [24, 25, 66-69]. PARP14 has also been shown to promote the Warburg effect in hepatocellular carcinoma by inhibiting JNK1-dependent PKM2 phosphorylation [70]. Such reports raise the possibility that targeting PARP14 may result in the development of effective therapies for inflammatory diseases and various cancers. In the future, it will be possible to conduct functional cell analyses using the putative inhibitors that I obtained, which may help to elucidate the contribution of PARP14 to these diseases. The compounds I identified have the potential to make a significant contribution to research concerning PARP14. In conclusion, my study provides an ideal platform for the discovery of PARP14 inhibitors, and presents new chemical tools for use in investigating the multiple functions of PARP14.

## Figures and Tables

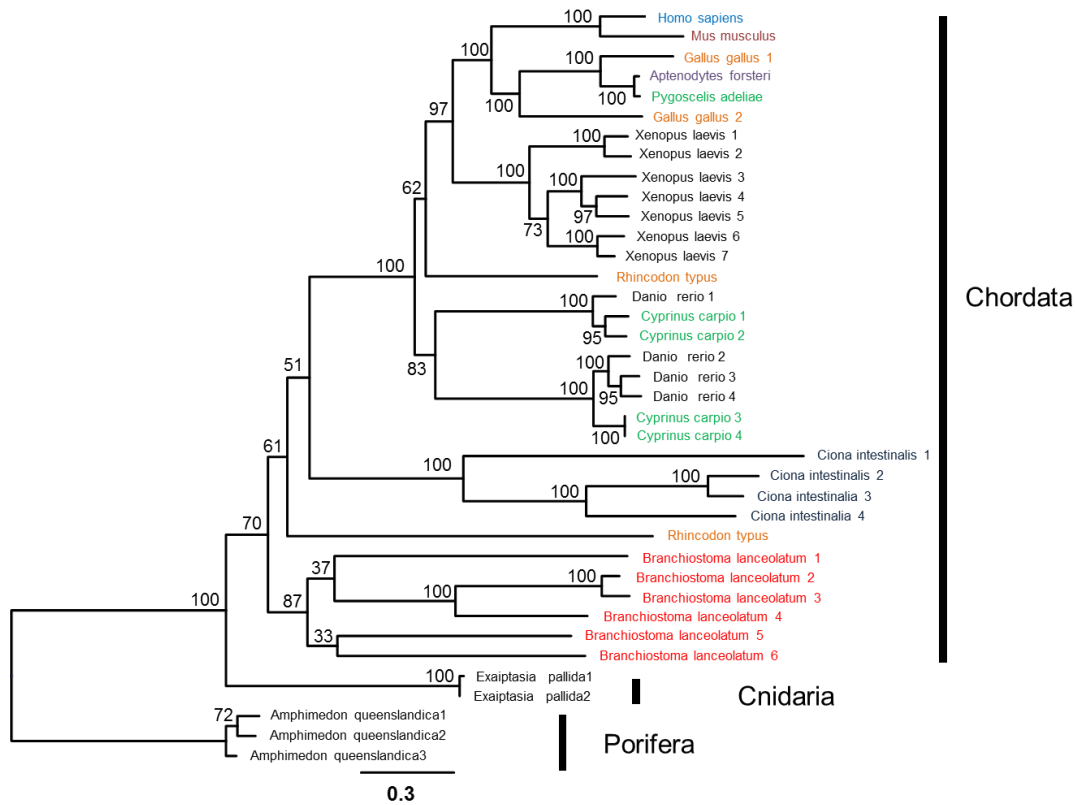


Figure 6. Phylogenetic tree of PARP14. The ML tree inferred by RAxML program using the GTR +  $\Gamma$  4 model was shown. Different colors were used to differentiate species in Chordata. Branch lengths are proportional to the number of amino acid substitutions/site. Numbers on the internal branches denote bootstrap proportion values.

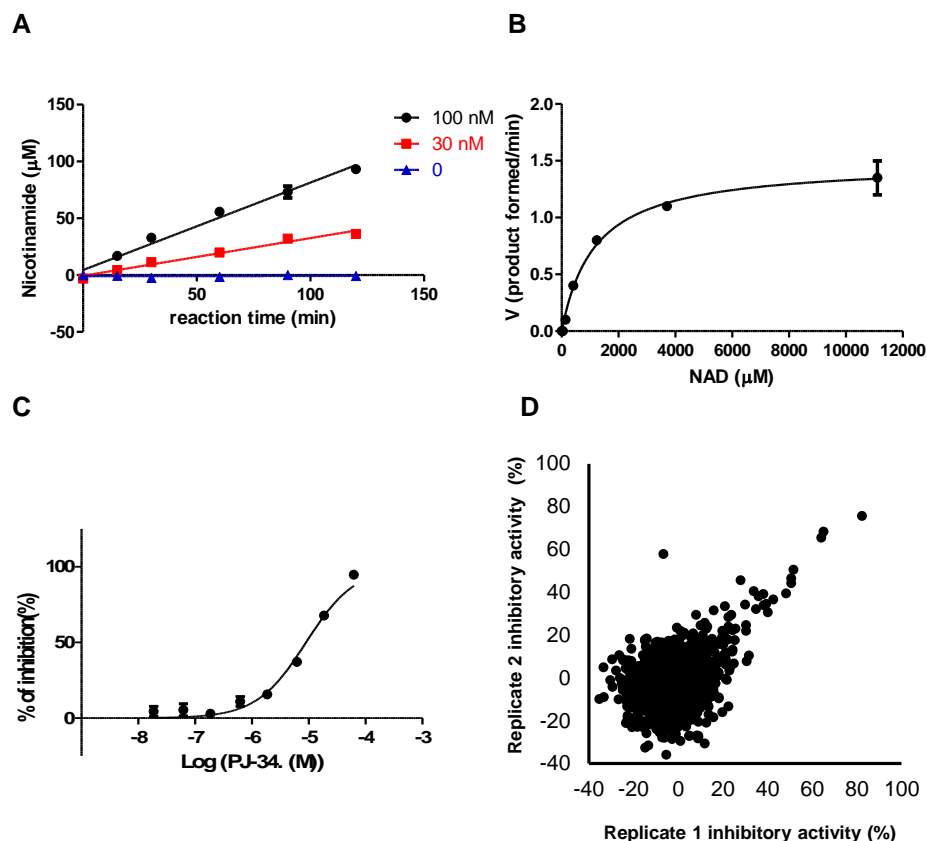


Figure 7. PARP14 high-throughput mass spectrometry (HTMS) assay development.

(A) Time course of PARP14 in the presence of 100 nM or 30 nM PARP14. Nicotinamide production by PARP14 increased in a linear manner for 120 min. (B) Determination of the kinetic constants for  $\text{NAD}^+$ . The study analysis was performed under the assay conditions of 0.1 μM PARP14 and an increasing concentration of  $\text{NAD}^+$ . The  $K_m$  value (95% confidence interval) for  $\text{NAD}^+$  was 1.2 mM (0.79–1.5 mM). (C) Dose-dependent inhibition of PARP14 by PJ-34. The  $\text{IC}_{50}$  (95% confidence interval) value was 9.0 μM (7.5–10 μM). (D) Correlation between replicates of percentage inhibition of the test compounds. *x*-axis is replicate 1 and *y*-axis is replicate 2 ( $r=0.34$ ). Data shown are mean  $\pm$  SD of  $n = 2$  technical replicates and are representative of two independent experiments.

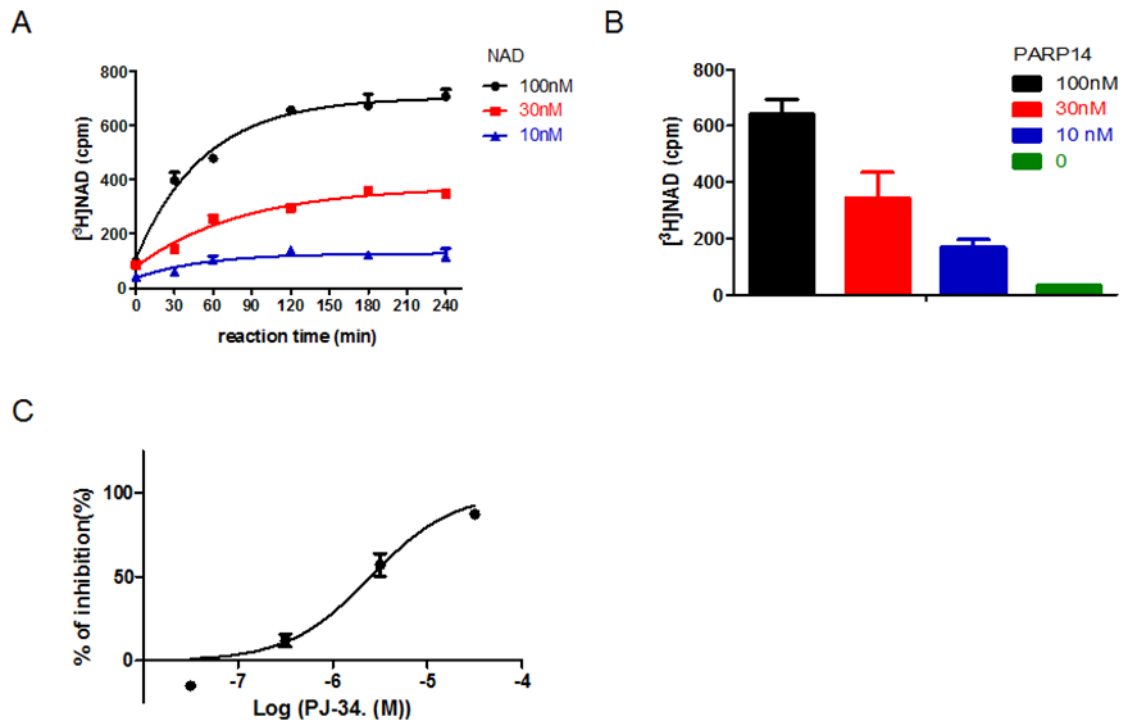


Figure 8. PARP14 immunoradiometric assay development.

(A) Time course of PARP14 auto-ribosylation activity at 100 nM, 30 nM, or 10 nM  $^3\text{H}$  NAD and 30 nM PARP14. (B) Auto-ribosylation activity with 100 nM, 30 nM, or 10 nM PARP14. (C) Dose-dependent inhibition of PARP14 by PJ-34.  $\text{IC}_{50}$  (95% confidence interval) value was  $2.5 \mu\text{M}$  ( $1.5\text{--}4.3 \mu\text{M}$ ). Data shown are mean  $\pm$  SD of  $n = 2$  technical replicates and are representative of two independent experiments.

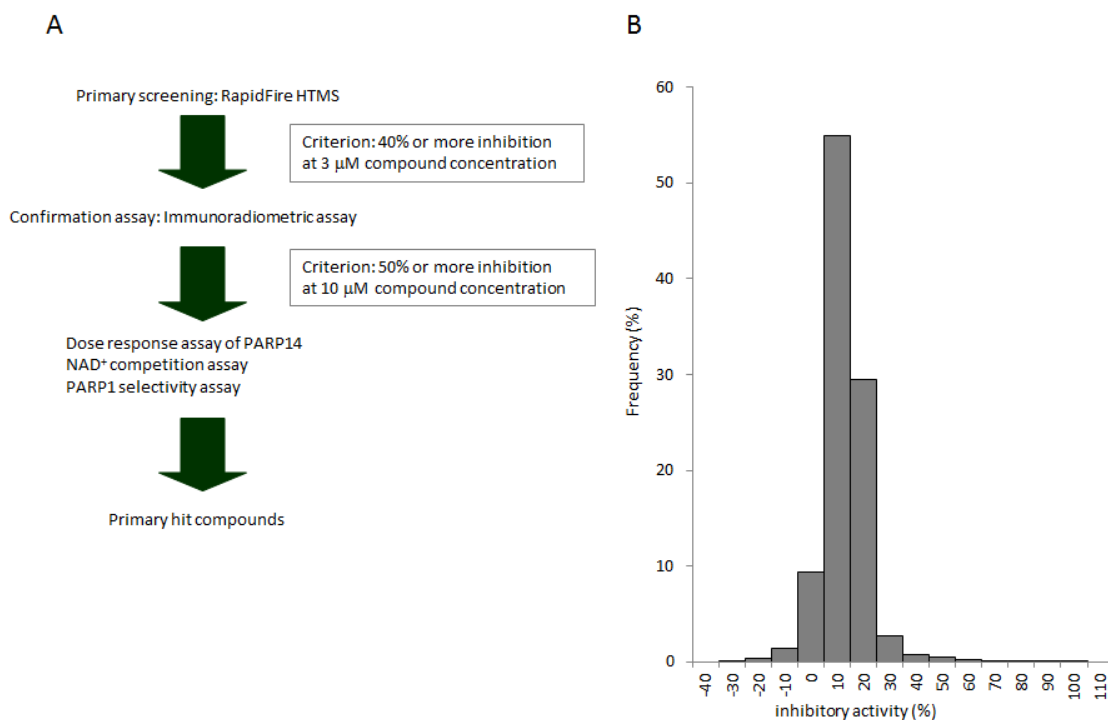


Figure 9. Results of primary screening and profiling assays.

(A) The HTS flow consists of primary screening, profiling assay, hit validation, and characterization. (B) A diverse library of small compounds at Takeda Pharmaceutical Company was screened at 3  $\mu$ M against PARP14, as shown by the histogram.

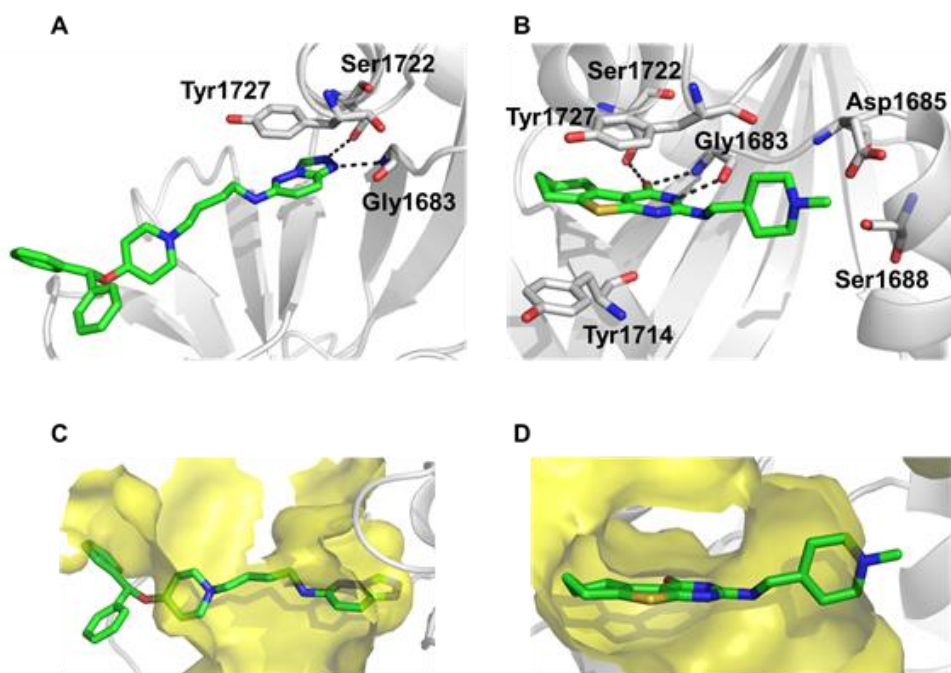


Figure 10 Crystal structures of PARP14 in complex with hit compounds.

(A, B) Interactions of ligands with PARP14 regarding compound 3 (A) and compound 4 (B). Hydrogen bonds between ligands and PARP14 are depicted as dotted lines. (C,D) Overlay of PARP1 and PARP14 in complex with compound 3 (C) and compound 4 (D), respectively. PARP1 protein is depicted as a yellow surface.

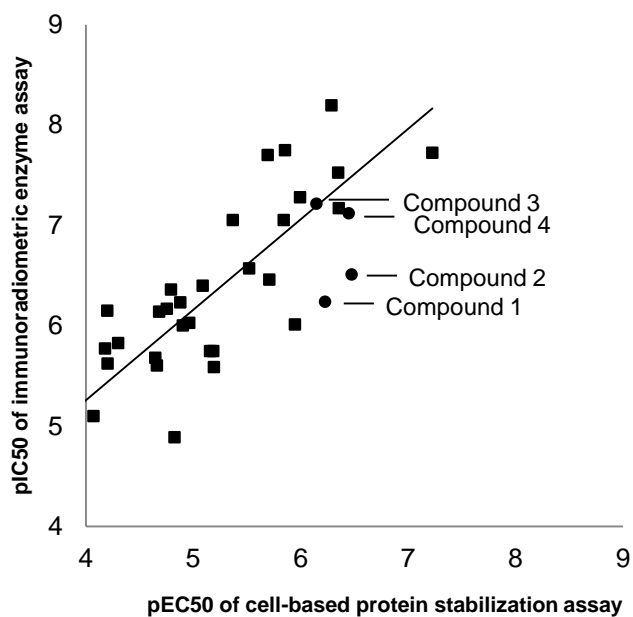
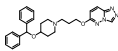
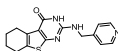


Figure 11. A correlation plot of the  $pIC_{50}$  values determined by enzyme immunoradiometric assay and cell-based protein stabilization assay ( $r=0.82$ ). The black circles represent compounds 1–4.

$IC_{50}$  values of compounds 1–4 are 590, 330, 710, and 350 nM, respectively.



Table 3. IC<sub>50</sub> values and 95% confidence intervals of the hit compounds in various assays

Compound ID	structure	Immunoradiometric assay		RapidFire HTMS			
		PARP14 IC <sub>50</sub> (μM)	PARP1 IC <sub>50</sub> (μM)	PARP14		PARP14	
				IC <sub>50</sub> (μM)		IC <sub>50</sub> (μM)	
				500 μM	NAD <sup>+</sup>	10 mM	NAD <sup>+</sup>
Compound 1		0.58 (0.38–0.87)	>30	2.8 (1.9–4.1)		13 (10–17)	
Compound 2		0.31 (0.27–0.36)	26 (5.5–91)	1.3 (1.1–1.6)		6.2 (4.1–9.3)	

**Table 4. X-ray data**

Data Collection	Compound 3		Compound 4	
Beamline	SSRL 7-1		SSRL 14-1	
Space group	P6 <sub>2</sub> 22		C222 <sub>1</sub>	
Unit cell parameters	a = 83.3 Å	α = 90°	a = 83.0 Å	α = 90°
	b = 83.3Å	β = 90°	b = 144.9Å	β = 90°
	c = 129.9 Å	γ = 120°	c = 83.4Å	γ = 90°
Wavelength (Å)	1.1271		1.2398	
Resolution range (Å)	50.0–2.35		30.0–2.6	
Number of observations	738,119		556,336	
Number of unique reflections	14,218		16,050	
Completeness (%)	95.5 (86.2)		99.8 (97.8)	
Redundancy	5.7		3.7	
Mean I/σ (I)	24.1 (1.9)		9.8 (1.2)	
Rsyma on I	0.079		0.171	
Highest-resolution shell (Å)	2.41–2.35		2.65–2.58	
Refinement				
R-work	22		20.3	
R-free	29.6		26.4	
Average B factor	43.1		23.6	
Rmsd bond length (Å)	0.005		0.007	
Rmsd bond angle (°)	0.889		1.05	
Ramachandran statistics				

Preferred region (%)	97.6	95.2
Allowed region (%)	1.8	4.18
Outlier (%)	0.6	0.6

---

Chapter 2 : High-throughput Screening to Identify  
Inhibitors of DEAD Box Helicase DDX41

## Chapter 2 : High-throughput Screening to Identify Inhibitors of DEAD Box Helicase DDX41

### Abstract

The human DEAD (Asp–Glu–Ala–Asp) box protein DDX41, a member of the DEXDc helicase family, has nucleic acid-dependent ATPase and RNA and DNA translocase and unwinding activities. DDX41 is affected by somatic mutations in sporadic cases of myeloid neoplasms as well as in a biallelic fashion in 50% of patients with germline DDX41 mutations. The R525H mutation in DDX41 is thought to play important roles in the development of hereditary myelodysplastic syndrome and acute myelocytic leukemia.

In this study, human DDX41 and its R525H mutant (R525H) were expressed in *Escherichia coli* and purified. The ATPase activities of the recombinant DDX41 and R525H proteins were dependent on both ATP and double-stranded DNA (dsDNA), such as poly(dG–dC) and poly(dA–dT). High-throughput screening was performed with a dsDNA-dependent ATPase assay using the human R525H proteins. After hit confirmation and counter screening, several small-molecule inhibitors were successfully identified. These compounds show DDX41-selective inhibitory activities.

## Introduction

A comprehensive genetic analysis identified a somatic mutation in the DDX41 gene encoding a DEAD (Asp–Glu–Ala–Asp) box-type adenosine triphosphate (ATP)–dependent RNA helicase. The somatic mutations of DDX41 are dominated by c.G1574A (p.R525H). Germline mutations in DDX41 were recently isolated in a subset of familial acute myeloid leukemia (AML)/myelodysplastic syndrome (MDS) pedigrees. Spliceosomal proteins constitute the top functional group associated with DDX41 and R525H alter the native DDX41 interaction especially for the U2 and U5 spliceosomes. It was considered that DDX41 defects lead to loss of tumor suppressor function due to altered pre-mRNA splicing and RNA processing [71]. The R525H mutation not only causes a defect in pre-rRNA processing but also inhibits cell cycle progression [72]. The R525H mutant (R525H) might also play an important role in tumorigenesis and therefore provides a drug target for cancer therapy [73-76]. R525H residue in DDX41 is located at the C-terminus of the RecA-like helicase core domain where ATP might interact. Though it is expected that R525H might show the different helicase activity compared to DDX41, the enzymatic activities are not evaluated so far.

DDX41 is a member of the DEXDc family of ATP dependent helicases. [77, 78] DEAD-box proteins, which contain the conserved sequence Asp–Glu–Ala–Asp, play essential roles in DNA and RNA metabolism, including in their replication, repair, recombination, transcription, translation, ribosome biogenesis, and splicing, which regulate growth and development [79, 80]. The analysis of phylogenetic tree revealed DDX41 is conserved in eukaryotes especially in animals [81]. Although DDX41 might be an important molecule in the progress of the STAT6 signal's evolution, the mechanism of how it functions is still unknown. DDX41 has also been shown to sense

microbial DNA, triggering the early induction of Chemokines and type I IFNs in mouse splenic myeloid dendritic cells [29, 33]. When it is stimulated with poly(dA–dT) or poly(dG–dC) in some cell-based assays, DDX41 combines with these polynucleotides physically and interacts with the downstream adaptor molecule STING and TBK1, triggering the activation of interferon regulatory factor 3 (IRF3) and nuclear factor-kappa B (NF-κB), which are important transcription factors in the regulation of type I IFN expression [34, 82]. A recent study also determined the crystal structures of the DDX41 DEAD domain in the apo form [73]. However, very little is known about the enzyme dynamics of DDX41 [72, 83]. In the present study, I cloned the full-length cDNA of human DDX41, R525H and ATPase defective E345Q mutant (E345Q), characterized the proteins' functions with an ATPase activity assay and used HTS to identify R525H inhibitors. E345Q which was expected to be defective in ATP hydrolysis activity based on an analysis of eIF4A3 [84], to confirm the difference in the ATPase activities of DDX41 and E345Q and rule out the possibility that the dsDNA-dependent ATPase activity observed with purified DDX41 was actually due to a contaminant. The purified human DDX41 and R525H enzymes showed ATPase activity against poly(dG–dC) or poly(dA–dT). On the other hand, E345Q showed no ATPase activity. Various compounds were found to inhibit R525H and DDX41. These results constitute significant advances in the biochemical research into DDX41

## Materials and Methods

### Cloning

The genes encoding DDX41, R525H, and E345Q (ATPase defective mutant) were amplified with polymerase chain reaction (PCR) using PrimeSTAR® Max DNA Polymerase (Takara Bio, Kusatsu, Japan), with which BamHI and NotI restriction sites were inserted upstream and downstream from DDX41, respectively. The primer sequences for DDX41 were F1, 5'-GTACTTTCAGGGATCCATGGAGGAGTCGGAACCCGA-3', and R1, 5'-TGCTCGAGTGCGGCCGCTCAGAAGTCCATGGAGCTGTGGG-3'. R525H and E345Q were constructed with overlap extension PCR. The first round of PCR to amplify R525H was performed with primers F1 and R2 (5'-TGTGTTTCCCGAGTGCCCGGTGCGGCCAAT-3'), and the open reading frame (Thermo Fisher Scientific, Waltham, MA) of the DDX41 gene was used as the template DNA. The second round of PCR was performed with primers F2 (5'-ATTGGCCGCACCGGGCACTCGGGAAACACA-3') and R1, and DDX41 was used as a template. In the third round of PCR, the products of the first- and second-round PCRs were used as the templates and amplified with the F1 and R1 primers. To construct E345Q, the first round of PCR was performed with primers F1 and R3 (5' GATCATGCGGTCAGCCTGGTCCAGGGCCAGGTAGCGACA-3'), and DDX41 was used as the template. The second round of PCR was performed with primers F3 (5'-CCCTGGACCAGGCTGACCGCATGATCGACATG-3') and R1, and DDX41 was used as the template. The third round of PCR was performed in the same way as that for R525H. The amplified gene was inserted into the pET21HH(V) vector using the In-Fusion HD Cloning Kit (Takara Bio) and introduced into *E. coli* JM109 to



isolate the plasmid. *dnaK* was amplified with PrimeSTAR GXL DNA Polymerase (Takara Bio) and restricted with the *NdeI* and *NheI* enzymes. The genomic DNA of *E. coli* BL21(DE3) was used as the template. The first round of PCR was performed with primers F4 (5'-AAGGAGATATACATATGGGTAAAATAATTGGTATCG-3') and R4 (5'-TCCCTGAAAGTACAGGTTCTCTTTTTTGTCTTTGACTTCTTCAAATTC-3'). The second round of PCR was performed with primers F5 (5'-AAGGAGATATACATATGGGTAAAATAATTGGTATCG-3') and R5 (5'-CGTTCAGTCCGCTAGCTCCCTGAAAGTACAGGTTCTC-3'), and the PCR product amplified with the F4 and R4 primers was used as the template. After its digestion with restriction enzymes, the amplified gene was inserted into the pET21HH/hTARS-Avi-His vector using the In-Fusion HD Cloning Kit (Takara Bio) and introduced into *E. coli* JM109 for plasmid isolation. The isolated recombinant plasmids carrying DDX41, R525H, E345Q, and *dnaK* were introduced into *E. coli* BL21(DE3) (Nippongene, Toyama, Japan).

#### Protein Expression and Purification

*E. coli* BL21(DE3) clones were grown at 37 °C in Luria–Bertani broth to an optical density of 0.4 at 600 nm (OD<sub>600</sub>). Isopropyl β-D -thiogalactoside was added at a final concentration of 1 mM and growth was continued for 5 h at 37 °C. The recombinant proteins were purified in two chromatographic steps. The cells were lysed with lysis buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 5 U/mL SEM nuclease, 20 mM imidazole, 10% [v/v] glycerol, and 1 mM dithiothreitol [DTT]). The lysate was sonicated three times for 15 s each and the suspension was centrifuged at 34,000 x g for 30 min at 4 °C. The supernatant was applied to an affinity NiNTA column and

equilibrated with AF buffer A (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, 10% [v/v] glycerol, and 1 mM DTT). The proteins bound to the column were eluted with a linear gradient of 0.005–0.5 M imidazole in AF buffer A. The positive fractions were collected and dialyzed against AF buffer B (AF buffer A containing 250 mM imidazole) at 4 °C. The dialyzed fractions were applied to a HiLoad 26/600 Superdex 200 pg (GE Healthcare UK Ltd., Little Chalfont, UK) previously equilibrated with GF buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 10% [v/v] glycerol, and 1 mM DTT). The fractions were pooled and eluted in GF buffer.

The presence of the recombinant protein in the collected fractions was checked by staining with Coomassie Brilliant Blue (Thermo Fisher Scientific) and immunoblotting with a DDX41-specific antibody (Abnova, Taipei, Taiwan), a His-tag-specific antibody (Wako Pure Chemical Industries, Ltd., Osaka, Japan) or a DnaK-specific antibody (in-house antibody).

### Mass Spectrometric Analysis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of the purified enzymes was performed with SYPRO Ruby Protein Gel Stain (Thermo Fisher Scientific). The protein bands were excised and digested with trypsin (Promega, Madison, WI) in a MultiScreen 96-well filter plate (Merck Millipore, Darmstadt, Germany). The tryptic digests were analyzed with nano-LC (Shimadzu, Kyoto, Japan) and an LTQ OrbitrapXL ETD Hybrid Ion Trap-Orbitrap Mass Spectrometer (Thermo Fisher Scientific), and identified using the Mascot database search system (Matrix Science, Boston, MA) and the SProt database. Data processing was performed by Proteome discoverer software ver. 1.4 (Thermo Fisher Scientific). The precursor mass

tolerance was set at 20 ppm, the fragment ion mass tolerance was 0.2 Da, and the false discovery rate was 0.01.

#### DDX41 Enzymatic Activity Assay

ATPase assays were performed with the ADP-Glo™ Kinase Assay (Promega), all of which were performed within the linear portion of the standard curve for the conversion of ATP to ADP. The kinetic parameters for ATP, poly(dG–dC), and poly(dA–dT) were determined by incubating various concentrations of each substrate with DDX41 or R525H for 30 min in 6 µL reaction volume in Greiner 784075 384-well microplates. The initial rate data were fitted to the Michaelis–Menten equation with GraphPad Prism v.5 (GraphPad Software Inc., San Diego, CA). The reactions for HTS were performed in a total volume of 1 µL in Greiner 782075 1536-well microplates and included 25 nM R525H enzyme, 70 µM ATP, and 0.2 µg/mL poly(dG–dC) (Sigma, Darmstadt, Germany) or poly(dA–dT) (Sigma) diluted with assay buffer (20 mM Tris-HCl [pH 7.5], 100 mM KCl, 2.5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.01% Tween 20, and 0.01% bovine serum albumin [Sigma]). After incubation for 120 min at room temperature, the reactions were terminated by the addition of 0.5 µL of ADP-Glo™ reagent (Promega). The samples were incubated for 40 min at room temperature before the addition of 1 µL of Kinase Detection Reagent™ (Promega). After 40 min at room temperature, the luminescence was read with an EnVision Multiplate Plate Reader (PerkinElmer, Waltham, MA). The assay quality of HTS was analyzed with the Z' factor, which was calculated from the means and standard deviations of the 100% and 0% inhibitory activities. I considered the luminescent signals of the reaction without poly(dG–dC) to indicate 100% inhibitory activity and those of the complete reaction mixture to indicate 0% inhibitory

activity. Curve fitting and calculation of the 50% inhibitory concentrations ( $IC_{50}$ ) were performed with GraphPad Prism v.5 (GraphPad Software).

The substrate competition assays using R525H enzyme were performed under the following four conditions: 70  $\mu$ M or 700  $\mu$ M ATP in the presence of 0.2  $\mu$ g/mL poly(dG–dC), and 0.2  $\mu$ g/mL or 2  $\mu$ g/mL poly(dG–dC) in the presence of 70  $\mu$ M ATP. The time-dependent inhibition was then confirmed by setting the preincubation between the R525H enzyme and the inhibitor to 0 min or 120 min before the reaction was initiated with 2  $\mu$ g/mL poly(dG–dC) and 70  $\mu$ M ATP. The reaction time was 5 min to confirm the change in the  $IC_{50}$  by setting different preincubation times of the enzyme with the inhibitor.

#### EIF4A3 Enzymatic Activity Assay

An RNA-dependent ATPase assay for eIF4A3 was also performed using the ADP-Glo™ assay system and proteins were prepared by the method described by Ito et al [84]. To enhance the ATPase activity of eIF4A3, 150 nM MLN51 was added to 150 nM eIF4A3 enzyme and 35  $\mu$ M ATP and 1.5  $\mu$ g/mL poly-U (MP Biomedicals, Santa Ana, CA) were used as the substrates. The reagents were incubated at room temperature for 30 min. Detection with ADP-Glo™ was performed as for DDX41. I considered the luminescent signals of the reaction without enzyme to indicate 100% inhibitory activity and those of the complete reaction mixture to indicate 0% inhibitory activity. Curve fitting and the calculation of the  $IC_{50}$  values were performed with GraphPad Prism v.5.

#### DnaK Enzymatic Activity Assay

ATPase activity of DnaK was also confirmed using the ADP-Glo™ assay system. The

reaction was performed with 24  $\mu\text{M}$  DnaK, 0.5  $\mu\text{M}$  ATP, and test compounds diluted with the same assay buffer as DDX41 at room temperature for 40 min. The detection with ADP-Glo™, the calculation of activity, and curve fitting were performed as for EIF4A3.

#### Fluorescent Intercalator Displacement Assay

Poly(dG–dC) (1  $\mu\text{g/mL}$ ) and 30  $\mu\text{M}$  candidate hit compound were incubated with 0.1  $\mu\text{M}$  TO-PRO®-3 stain (Thermo Fisher Scientific) in DDX41 assay buffer for 120 min in Greiner 784075 384-well microplates. The displacement activities were detected as the fluorescence at an excitation wavelength of 512 nm and an emission wavelength of 533 nm. I considered the fluorescent signals of the reaction without TO-PRO-3 to indicate 100% intercalating activity and those of the complete reaction mixture to indicate 0% intercalating activity.

#### HPLC analysis

Analytical HPLC was conducted using a corona charged aerosol detector (Thermo Fisher Scientific, Waltham, M.A.). An L-column 2 ODS [ $30 \times 2.1$  mm I.D., (CERI, Tokyo, Japan)] was used at a temperature of 50 °C and a flow rate of 0.5 mL/min. Mobile phases A and B under a neutral condition comprised a mixture of 50 mmol/L AcONH<sub>4</sub>, water, and MeCN (1/8/1, v/v/v) and a mixture of 50 mmol/L AcONH<sub>4</sub> and MeCN (1/9, v/v), respectively. The ratio of mobile phase B was increased linearly from 5% to 95% over 3 min and 95% over the next 1 min.

### X-ray structure analysis of the compounds

All measurements were performed on a Rigaku XtaLAB P200 diffractometer using graphite monochromated Cu-K $\alpha$  radiation. The structure was solved by direct methods with software SIR2008 and was refined using full-matrix least-squares on  $F^2$  with SHELXL-2013/4 [85]. All non-hydrogen atoms were refined with anisotropic displacement parameters. Full crystallographic data have been deposited with the Cambridge Crystallographic Data Centre (reference numbers: CCDC 1535690 for compound A, CCDC 1535691 for compound B). Copies of the data can be obtained free of charge via the following URL: <http://www.ccdc.cam.ac.uk/Community/Requestastructure/Pages/DataRequest.aspx?>

## Results

### Activities of DDX41 and Mutants

The recombinant DDX41, R525H, and E345Q proteins were expressed in *E. coli* and purified through NiNTA and Superdex 200 columns. On SDS-PAGE and western blotting, the mutant proteins appeared around the DDX41 band at 83 kDa (Fig. 12A, B). A protein mass analysis confirmed the types of proteins contained in the purified DDX41, R525H, and E345Q bands. A protein mass analysis revealed that the MS-identified peptides covered human DDX41 (Table 5), and DnaK from *E. coli* was one of the proteins included with the purified DDX41 protein (Table 6). A western blotting analysis of the purified DDX41 protein detected the DnaK protein with a DnaK-specific antibody (Fig. 12C).

The results of the ATPase activity assay are shown in Fig. 13 and Fig. 14. Both DDX41 and R525H displayed double-stranded DNA (dsDNA)-dependent ATPase activity, whereas E345Q did not. Analysis of the substrate specificity of DnaK revealed that DnaK did not use dsDNA as a substrate. I defined the luminescent signal of the reaction without dsDNA as indicating 100% inhibitory activity and that of the complete reaction mixture as indicating 0% inhibitory activity, and measured the dsDNA-dependent activities of DDX41 and R525H accordingly.

The apparent  $K_m$  values for dsDNA and ATP were determined according to the Michaelis–Menten equation, and their values are shown in Fig. 14 and Tables 7 and 8, respectively. DDX41 and R525H showed the same affinities for both substrates. The substrate concentrations under the screening conditions were 70  $\mu$ M ATP and 0.2  $\mu$ g/mL poly(dG–dC). The concentrations of the substrates were close to the apparent  $K_m$  values to maximize the diversity of screening hits [86] and to maintain the assay quality

throughout the screening process.

#### Identification of Novel R525H Inhibitors with HTS

A diverse library of more than 500,000 small-molecule compounds (10-compound mixture-based library) from the Takeda Pharmaceutical Company Limited was screened at a concentration of 3  $\mu$ M against R525H, as shown in the work flow (Fig. 15A). The assay quality was deemed acceptable because the  $Z'$  factor was 0.78 and the signal-to-noise ratio was 6.1. The histogram of the primary screening in Fig. 15B shows a normal distribution curve with a median value of  $-0.69$  and a standard deviation (SD) of 7.8. The hit frequency was 0.5% when the threshold was set at  $\geq 31\%$  of the median + 3SD inhibitory activity. These primary hit compounds through a process of deconvolution were picked and their inhibitory potency for EIF4A3 was tested with the ADP-Glo™ assay to eliminate the nonselective compounds. In this step, 12% of the initial hits were eliminated and 180 compounds with no activity against eIF4A3 were selected (Fig. 16A). All compounds showed the same inhibitory activities for DDX41 and R525H (Fig. 16B) and none of the compounds showed inhibitory activity for DnaK (Fig. 16C). Moreover, in-house database was searched to obtain the compounds which showed no activity against other target molecules such as kinases, G protein-coupled receptors, and other enzymes and 43 compounds were selected. The presence of false positive compounds that interacted as intercalators with poly(dG–dC) was a concern, so an intercalator displacement assay using TO-PRO-3 was performed to remove such compounds. The threshold was set at 29% of the mean of the low control + 3SD of the low control and a total of 13 of the 43 compounds showed displacement activity (Fig. 16D). The assay to remove the compounds quenching of TO-PRO-3 fluorescence was



also performed under the condition without poly(dG-dC) and 4 compounds were eliminate. Ultimately, I identified 26 hit candidate compounds, such as compound A [ (4E)-N-hydroxy-1-phenyl-2,3-dihydroquinolin-4(1H)-imine, CAS Registry Number: 10258-02-3] and compound B (4-amino-6-((2-amino-1,6-dimethylpyrimidin-4(1H)-ylidene)amino)-1-methyl-2-phenyl quinolinium chloride hydrochloride or the tautomer, CAS Registry Number: 115230-55-2), that interacted selectively with DDX41 and R525H from the results described above (Fig. 16E). The structures were confirmed by single crystal X-ray analysis (Fig. 17,18) and the purities of the compounds were shown to be 95% or more by LC-MS.

#### Characterization of the Hit Compounds

To understand the mechanism of inhibition, substrate competition assays were performed. The rates at which the compounds inhibited R525H were examined under the four conditions described in the Materials and Methods. Representative hit compounds, exemplified by compound A and compound B, showed the theoretical shift in their  $IC_{50}$  values when the poly(dG-dC) concentration was increased (Fig. 19A, B). However, the  $IC_{50}$  values for both compounds shifted only slightly when the ATP concentration was increased. These results indicate that these hit compounds may be competitive inhibitors of DDX41 with respect to the poly(dG-dC) substrate and non-competitive inhibitors with respect to the ATP substrate (Fig. 19C, D). Although compounds such as DNA intercalators were ruled out by the TO-PRO-3 competition assay, there may be other modes of DNA binding besides intercalation, such as minor groove binding, major groove binding and electrostatic binding to the backbone. Further

analyses would be necessary to draw definitive conclusions on the inhibitory mechanism. When the preincubation time (before the reaction was initiated by the addition of the substrates) for the enzyme and the test compound was set at 0 min or 120 min, the inhibitory activities of the compounds remained the same and it was confirmed that the compounds are not time-dependent inhibitors (Fig. 19E, F).

## Discussion

The results show that the ATPase activities of human DDX41 and R525H are dependent on ATP and dsDNA, either poly(dG–dC) or poly(dA–dT). Novel inhibitors of DDX41 and R525H were identified with HTS. Two hit compounds identified with several assays showed selective inhibitory activities.

DnaK was present in the purified DDX41 and R525H enzymes, detected with mass spectrometry and western blotting. The DnaK protein is a molecular chaperone and folding catalyst, with ATPase activity in the presence of ATP [87]. I tried to prepare high-purity enzymes by removing the contaminated proteins, such as DnaK, with anion exchange column chromatography. However, this caused the aggregation of DDX41. It is possible that DnaK assists in the folding of DDX41 in *E. coli*. The activities of DDX41/R525H and DnaK could be differentiated based on their dsDNA-dependent ATPase activities because the human E345Q ATPase defective mutant and DnaK showed no dsDNA-dependent ATPase activity. Therefore, I set assay conditions in the presence or absence of dsDNA to represent the maximum and minimum controls to detect the activities of DDX41 and R525H.

DDX41 may have both ATPase and helicase activities, but no assay for helicase activity has yet been reported. The  $K_m$  values for dsDNA against DDX41 and R525H were very similar. I speculate that these values would differ if the enzymes were analyzed with helicase activity assays. Pause et al [88] reported that mutants R365Q and R365K of eIF4A3, whose position corresponds to the conserved residue R525 in DDX41, exhibited dramatic reductions in the helicase activity for dsRNA, but did not show a significant difference in ATPase activity. The poly(dG–dC) and poly(dA–dT) used in this experiment were purchased from Sigma, and the lengths of these dsDNAs

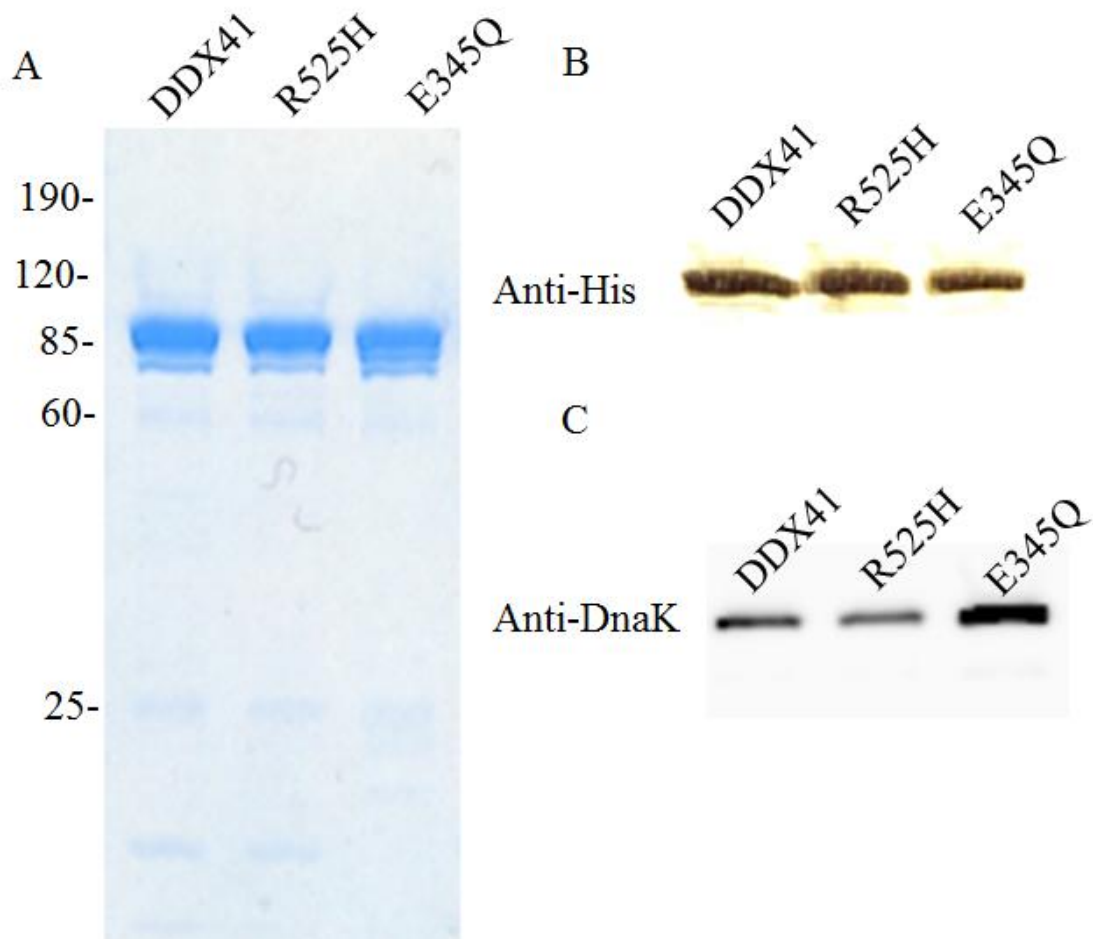
were not accurately characterized. If the sequence of the dsDNA or dsRNA recognized by DDX41 is determined and a helicase assay is established, the hypothesis that DDX41 and R525H have different helicase activities can be tested [89].

Several compounds were identified with HTS. In parallel with HTS, I developed a selectivity assay to verify the specificity of the enzymatic inhibition. The hit compounds, compound A and compound B, displayed inhibitory activities with IC<sub>50</sub> values of 4.3  $\mu$ M and 1.5  $\mu$ M, respectively, against human R525H mutant under the HTS conditions used and no activity for eIF4A3 and DnaK. Although I initially selected EIF4A3 because of its ready availability, the further selectivity analysis is necessary using another helicase with more similar sequences with DDX41. In general, the DNA- or RNA-unwinding activities of helicases are dependent on their ATPase activity. Therefore, these hit compounds may also inhibit the helicase activity of DDX41 and R525H. HTS results suggest that DDX41 is a potential drug target because no inhibitors of DDX41 have yet been identified. There may be a prominent binding site for small-molecule inhibitors in DDX41 that overlaps the poly(dG–dC)-binding site. The crystal structures of the DDX41-bound inhibitors could clarify their mode of inhibition and facilitate the optimization of lead compounds.

It has been reported that DDX41 interacts directly with several proteins including STING and TBK1, which is a crucial adaptor molecule for cytoplasmic DNA receptors [33]. The silencing of DDX41 expression in murine myeloid dendritic cells led to a marked reduction in IFN and proinflammatory cytokine production in response to cytosolic DNA [83]. Besides its role in immunity, the human R525H mutation is associated with familial MDS and AML. Xenograft experiments with cell lines in which DDX41 was knocked down demonstrated DDX41-enhanced tumor growth, but the

mechanism behind this is still not understood. I have identified inhibitors of both DDX41 and R525H, with which it should be possible to conduct biochemical analyses. The clarification of their efficacy in tumor cells, the molecular mechanism of DDX41, and the crystal structure of DDX41 when bound to these compounds will have a significant impact on research into DDX41. If it is possible to create a DDX41 or R525H selective inhibitor by lead generation, the significance of the mutation could be clarified.

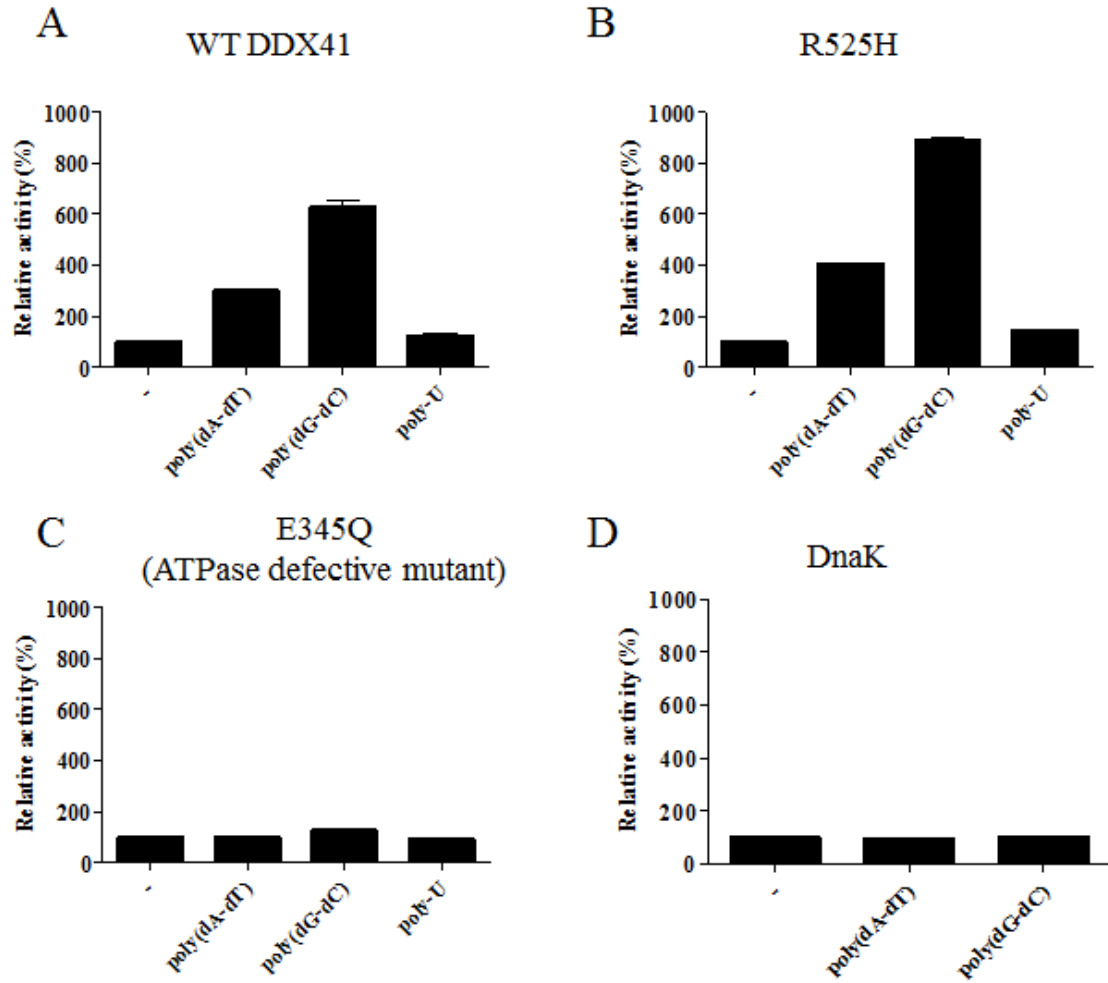
In conclusion, my study provides evidence that DDX41 has dsDNA-dependent ATPase activity, dependent on poly(dG–dC) or poly(dA–dT), and presents new chemical tools for further investigation of the multiple functions of DDX41.



**Figure 12.** Analysis of purified enzymes.

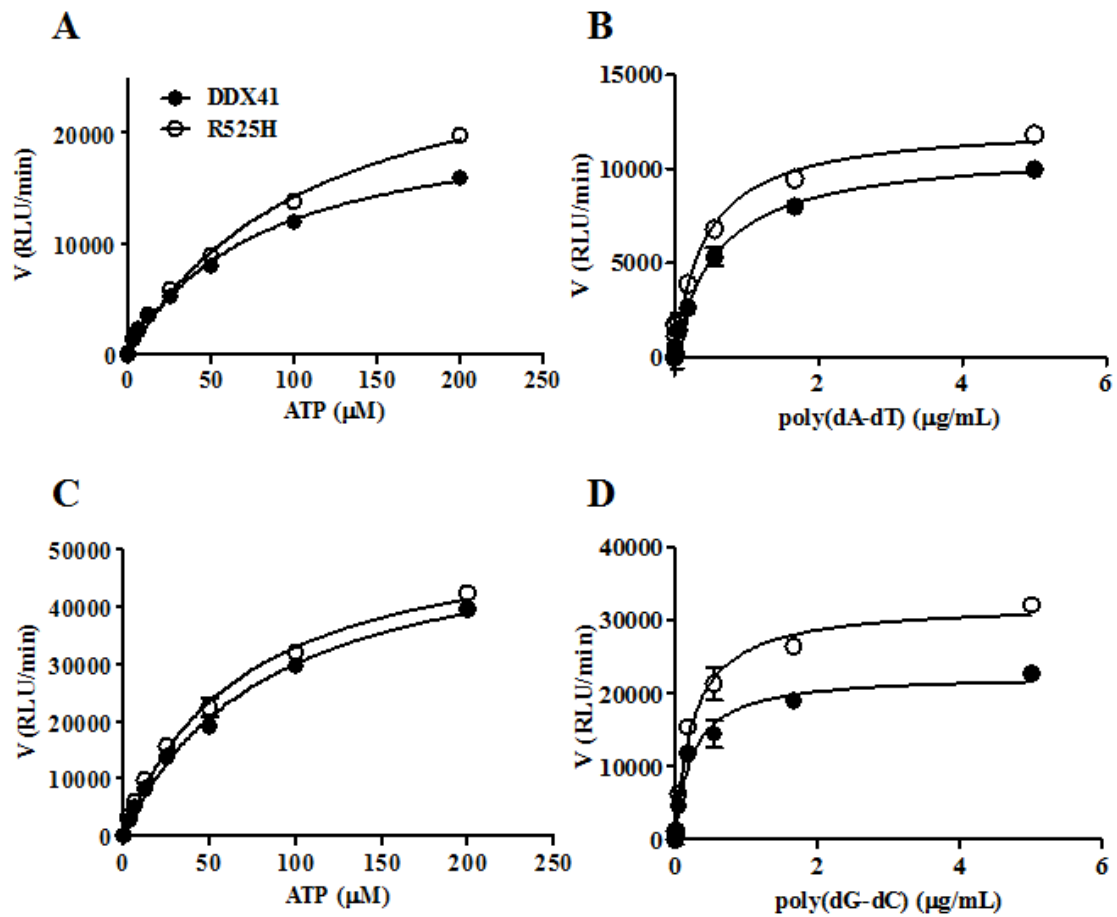
SDS-PAGE of DDX41, R525H, and E345Q stained with Coomassie Blue (A) and a western blot probed with an anti-His-tag antibody (B) or anti-DnaK antibody (C).

DnaK from *E. coli* was one of the proteins retained in the purified DDX41.



**Figure 13.** Substrate specificities of DDX41, R525H, E345Q, and DnaK.

The substrate specificities were confirmed with 5  $\mu\text{g/mL}$  poly(dA-dT), poly(dG-dC), and poly-U in the presence of 100  $\mu\text{M}$  ATP and 500 nM enzyme. DDX41 (A) and R525H (B) showed poly(dA-dT)- or poly(dG-dC)-dependent activities, but E345Q (C) and DnaK (D) were not activated by dsDNA. Relative activity values were calculated by setting 100% as the activity without dsDNA or dsRNA. Data shown are mean  $\pm$  SD of  $n = 2$  technical replicates and are representative of two independent experiments. WT, wild type

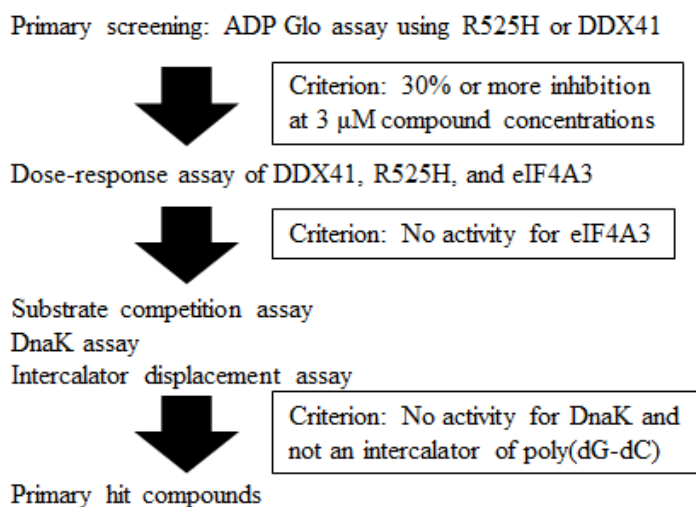


**Figure 14.** Determination of the kinetic constants for ATP and dsDNA (poly[dA–dT] and poly[dG–dC]) in ATPase reactions with recombinant DDX41 or R525H protein.

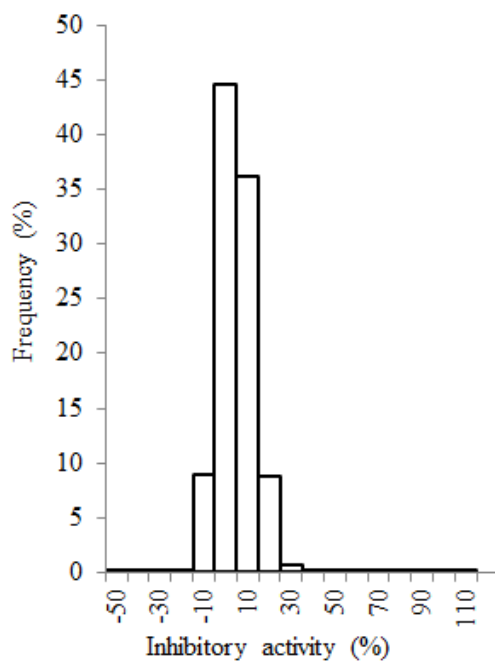
The assay was performed with 5  $\mu$ g/mL dsDNA and increasing concentrations of ATP (A) and (C); or 100  $\mu$ M ATP and increasing concentrations of dsDNA (B) and (D).  $K_m$  values for poly(dA–dT) and poly(dG–dC) with ATP are shown in Tables 7 and 8, respectively. RLU, relative luminescence units. Data shown are mean  $\pm$  SD of  $n = 2$  technical replicates and are representative of two independent experiments.



**A**

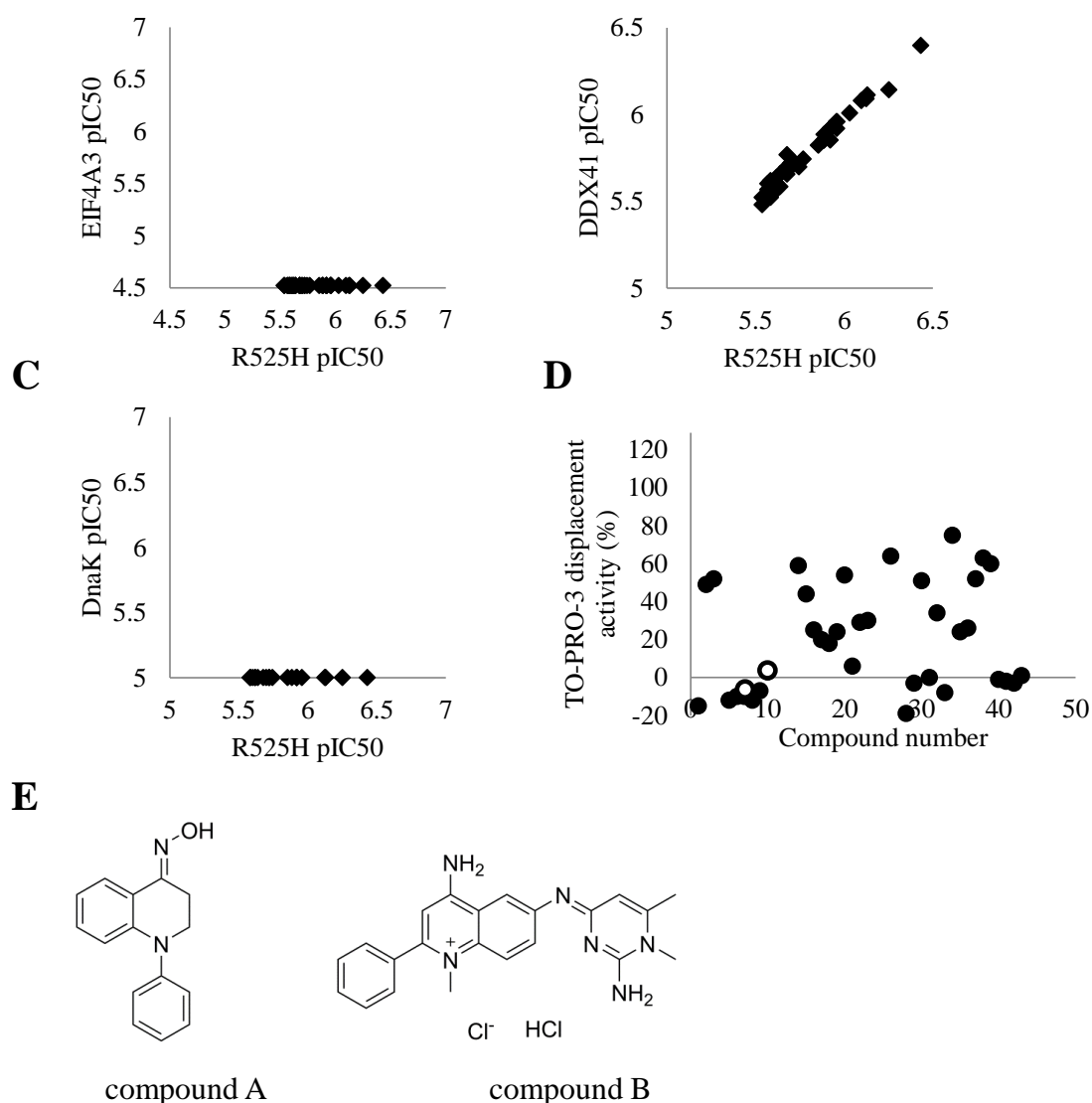


**B**



**Figure 15.** Screening cascade and results of primary screening.

The HTS cascade consisted of the primary screening, profiling assay, hit validation, and characterization (A). A diverse library of small compounds from Takeda Pharmaceutical Company at a concentration of 3  $\mu$ M was screened against DDX41, as shown in the histogram (B). The hit frequency of the primary screening was 0.5%; hits showed  $\geq 31\%$  of the median + 3SD inhibitory activity.



**Figure 16.** Results of profiling assays

The primary hit compounds were analyzed with a selectivity assay using eIF4A3 (A), DDX41 (B), and DnaK (C). An intercalator displacement assay using TO-PRO-3 was conducted to remove compounds that interacted as an intercalator with poly(dG–dC) (D). The two white circles represent compound A and compound B. The structures of compounds A and B are shown in (E).

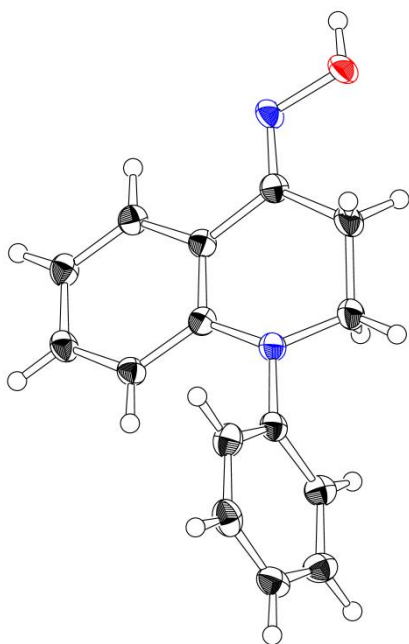


Figure 17. ORTEP of compound **A**; thermal ellipsoids are drawn at 50% probability.

Crystal data for (4*E*)-*N*-hydroxy-1-phenyl-2,3-dihydroquinolin-4(1*H*)-imine (compound **A**): C<sub>15</sub>H<sub>14</sub>N<sub>2</sub>O, MW = 238.29; crystal size, 0.18 × 0.17 × 0.06 mm; colorless, block; monoclinic, space group *C*2/*c*, *a* = 17.3229(9) Å, *b* = 8.0063(3) Å, *c* = 18.7849(9) Å,  $\alpha = \gamma = 90^\circ$ ,  $\beta = 107.059(6)^\circ$ , *V* = 2490.7(2) Å<sup>3</sup>, *Z* = 8, *D*<sub>x</sub> = 1.271 g/cm<sup>3</sup>, *T* = 100 K,  $\mu = 0.646 \text{ mm}^{-1}$ ,  $\lambda = 1.54184 \text{ Å}$ , *R*<sub>1</sub> = 0.051, *wR*<sub>2</sub> = 0.154.

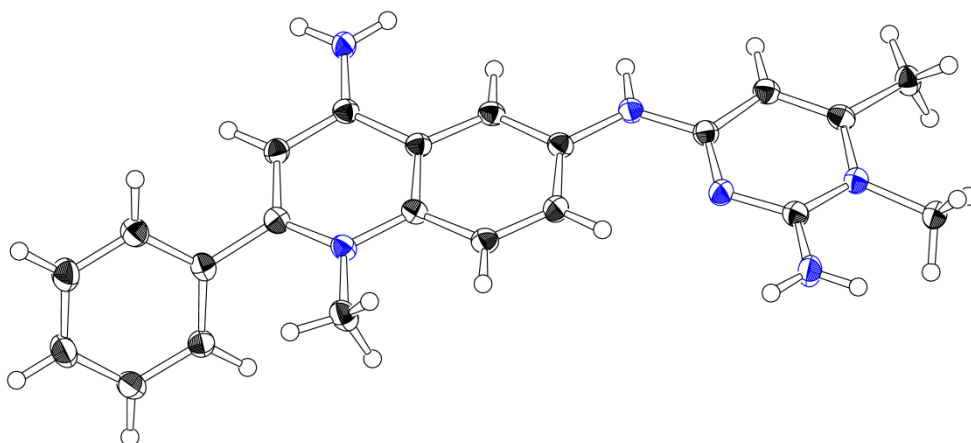
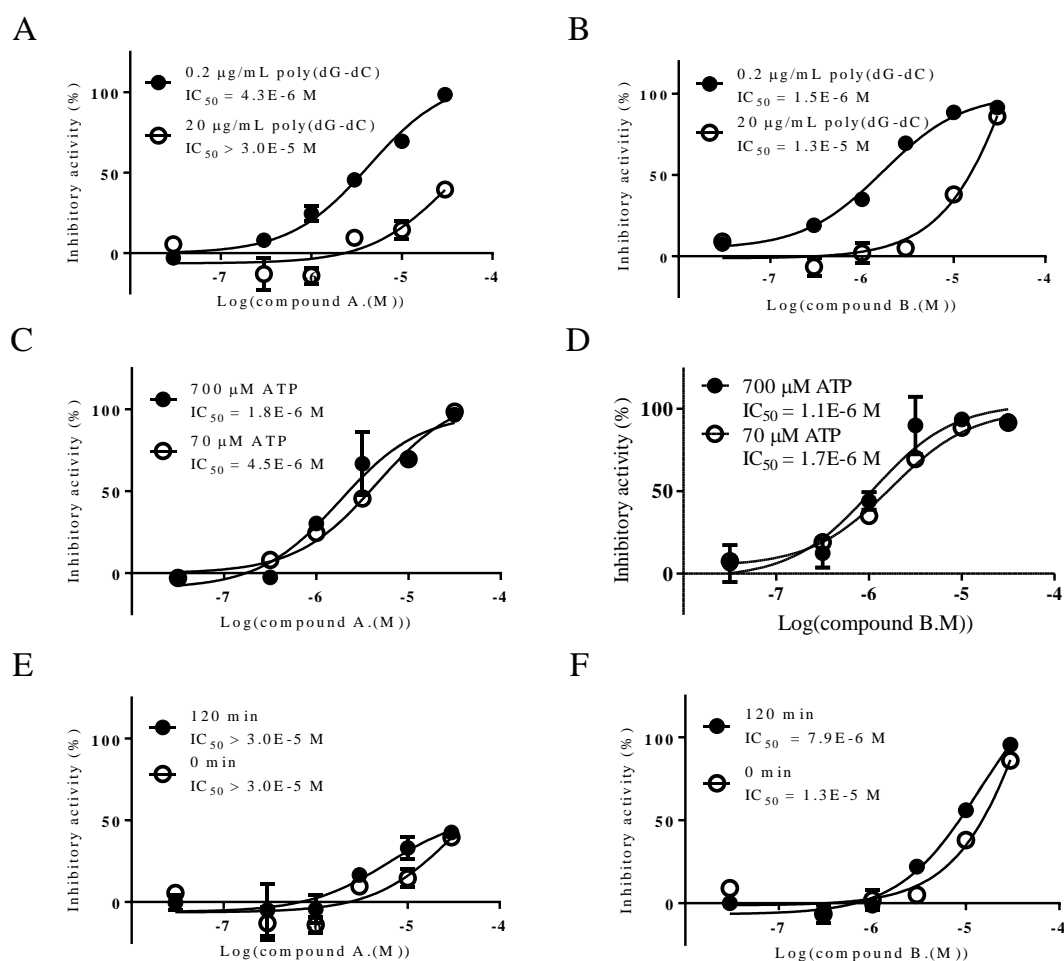


Figure 18. ORTEP of compound **B**; thermal ellipsoids are drawn at 50% probability.

Crystal data for 4-amino-6-((2-amino-1,6-dimethylpyrimidin-4(1*H*)-ylidene)amino)-1-methyl-2-phenylquinolinium chloride hydrochloride or the tautomer (compound **B**):  $\text{C}_{22}\text{H}_{24}\text{N}_6^{2+} \cdot 2\text{Cl}^- \cdot \text{H}_2\text{O} \cdot \text{C}_2\text{H}_6\text{O}$ , MW = 507.46; crystal size,  $0.10 \times 0.05 \times 0.04$  mm; colorless, block; triclinic, space group *P*-1,  $a = 10.8559(2)$  Å,  $b = 11.6852(3)$  Å,  $c = 11.8228(4)$  Å,  $\alpha = 118.093(3)^\circ$ ,  $\beta = 101.204(2)^\circ$ ,  $\gamma = 98.906(2)^\circ$ ,  $V = 1244.17(7)$  Å<sup>3</sup>,  $Z = 2$ ,  $D_x = 1.354$  g/cm<sup>3</sup>,  $T = 100$  K,  $\mu = 2.622$  mm<sup>-1</sup>,  $\lambda = 1.54184$  Å,  $R_1 = 0.040$ ,  $wR_2 = 0.107$ . Compound **B** was crystallized in 1:1:1 complex with H<sub>2</sub>O and ethanol.



**Figure 19.** Profiling assays of compound A and compound B with R525H.

The results for compound A are shown in (A), (C), and (E) and those for compound B are shown in (B), (D), and (F). Dose–response curves in the presence of 0.2 µg/mL or 20 µg/mL poly(dG–dC) were determined (A) and (B); the  $IC_{50}$  values of these compounds were increased by a high concentration of poly(dG–dC). Dose–response curves in the presence of 70 µM or 700 µM ATP were determined (C) and (D); the  $IC_{50}$  values of these compounds were not changed by a high concentration of ATP. Dose–response curves were determined by setting the preincubation period for the enzyme and inhibitor at 0 min or 120 min before the reaction was initiated with 20 µg/mL poly(dG–dC) and 70 µM ATP, to confirm the time-dependent inhibition (E) and (F). The x-axes indicate the concentration of the compounds and the y-axes indicate % inhibition. Data shown are mean  $\pm$  SD of  $n = 2$  technical replicates and are representative of two independent experiments.



Table 5. Human proteins identified by mass spectrometric analysis

Accession	Description	Sequence coverage	Number of matched peptides	MW (kDa)
Q9UJV9	Probable ATP-dependent RNA helicase DDX41	48.07	32	69.8
P04264	Keratin,type IIcytoskeletal 1	12.42	7	66.0
P13645	Keratin,type Icytoskeletal 10	6.34	3	58.8

Table 6. *Escherichia coli* proteins identified using mass spectrometric analysis

Accession	Description	Sequence coverage	Number of matched peptides	MW [kDa]
A7ZHA4	Chaperone protein DnaK	65.7	34	69.1
P39451	Alcohol dehydrogenase, propanol-preferring	55.1	11	35.4
P17169	Glucosamine-fructose-6-phosphate aminotransferase	51.0	22	66.9
P35340	Alkyl hydroperoxide reductase subunit F	21.9	8	56.1
A7ZTY0	Guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase	17.4	7	54.9
B7MG22	Bifunctional polymyxin resistance protein ArnA	14.9	7	74.3

Table 7. Apparent  $K_m$  values of ATP and poly(dG-dC)<sup>a</sup>



Enzyme	ATP apparent $K_m$ ( $\mu$ M)	Poly(dG-dC) apparent $K_m$ ( $\mu$ g/mL)
DDX41	79 (61–96)	0.59 (0.46–0.71)
R525H	113 (92–133)	0.42 (0.26–0.58)

<sup>a</sup>The numbers in parentheses represent 95% confidence intervals.

Table 8. Apparent  $K_m$  values of ATP and poly(dA-dT)<sup>a</sup>

Enzyme	ATP apparent $K_m$ ( $\mu$ M)	Poly(dA-dT) apparent $K_m$ ( $\mu$ g/mL)
DDX41	82 (66–97)	0.24 (0.15–0.32)
R525H	66 (54–79)	0.25 (0.19–0.32)

<sup>a</sup>The numbers in parentheses represent 95% confidence intervals.

## General Discussion

## General discussion

Drug discovery research can generally be conducted by the flow shown in Fig. 20. Developing a new drug from an original idea to its launch is a complex process that can take up to 15 years [82, 83]. Fig. 20 indicates the key preclinical stages of the drug discovery process from initial target identification and validation, through assay development, high throughput screening, hit identification, lead optimization, and finally, the selection of a candidate molecule for clinical development.

My research described in this thesis is positioned at the first stage. Generally, in the first stage, molecular, biological, and biochemical analyses are conducted to establish the basic knowledge to discover new drugs. The enzyme molecular dynamics can be evaluated successfully with some original assay methods and selective inhibitors are obtained through HTS. In my research, novel assay methods were established to detect PARP14 and DDX41 enzymatic activities. In the research on PARP14, I described the identification of potent, selective, and cell-permeable inhibitors through HTS. A strategy for creating potential lead compounds could be developed by solving the X-ray co-crystal structure of PARP14 with the hit compounds and SAR information. In the research on DDX41, the ATPase activities of human DDX41 and R525H were shown to be dependent on ATP and dsDNA, either poly(dG–dC) or poly(dA–dT). Novel inhibitors of DDX41 and R525H were identified by HTS. The hit compounds identified using several assays showed selective inhibitory activities. Therefore, my research provides new evidences on the properties of the STAT6 signal-related enzymes, PARP14 and DDX41, and presents chemical tools for further investigation of the multiple functions of both targets.

The STAT6 pathway is generally well-characterized as it has been implicated in related

to multiple diseases such as inflammation and cancer. There are two primary STAT6-mediated cellular pathways, i.e. the canonical and noncanonical pathways. The canonical pathway is activated by ligand interaction with IL-4/IL-12 receptors. Signal transduction of the canonical pathway results in the expression of GATA3. The noncanonical pathway uses a largely different set of proteins within its signaling cascade from those of the canonical one. A specific set of proteins including DDX41, STING, and TBK1, are responsible for inducing the noncanonical signaling. Transduction of this alternative pathway is dependent on the kinds of immune cells and ultimately results in cytokine production. Despite extensive validation of these two branches of the STAT6 pathway and their associated targets, considerable efforts within the pharmaceutical industry and academia have failed to produce therapeutics with clinical efficacy. These efforts have primarily adopted target-driven approaches, focusing mainly on creating inhibitors against STAT6 dimers and JAK because of their central functions within the pathway in the context of immunity. Although a number of these inhibitors have been developed, none so far have demonstrated therapeutic utility in the clinic. It has been difficult to obtain inhibitors with high potency and selectivity against these molecules. Therefore, I considered that PARP14 and DDX41 have possibilities as novel drug targets.

Based on the results of my enzymological research using molecular biological and biochemical strategies, the phase of drug discovery research for PARP14 and DDX41 will move to the lead generation stage and the role of these molecules will be clarified using the chemical probes that I obtained.

Lead compounds that possess higher inhibitory activities against PARP14 and appropriate pharmacokinetics might be synthesized based on the information from

X-ray co-crystal structures with hit compounds in the future. Moreover, the expression differences of target genes such as *Gata3* and cytokines secreted by Th2 cells and B cells can be evaluated using the hit compounds to determine how PARP14 contributes to STAT6 signaling. The effects of these compounds on the differentiation of Th2 cells and B cells were also confirmed. In addition to allergy, PARP14 has been shown to promote the Warburg effect in hepatocellular carcinoma by inhibiting JNK1-dependent PKM2 phosphorylation [62]. Such reports raise the possibility that targeting PARP14 may result in the development of effective therapies for inflammatory diseases as well as various cancers.

In the case of DDX41, selective inhibitors of R525H might be created first. It might be effective to obtain the crystal structure of both WT and R525H because it provides the information to determine the structural differences affecting the compound selectivity. If I successfully create selective compounds for R525H mutants, experiments analyzing the expression of cytokines such as CCL2 and CCL20 will be conducted with or without the tool compounds. As already mentioned in this thesis, the human R525H mutation is associated with familial MDS and AML. Xenograft experiments with cell lines in which DDX41 is knocked down could be performed to demonstrate DDX41-enhanced tumor growth. In order to uncover the underlying mechanism, the efficacy of the target could be clarified with tumor cells using R525H selective inhibitors.

This study clarified that PARP14 was present in the common ancestor of all metazoa. On the other hand, DDX41 was conserved in metazoa, plants, and fungi. Both molecules might play important roles within the organisms from the perspective of the cellular evolutionary process. The tool compounds might contribute to analyze these

molecules more deeply.

In summary, it will be possible to conduct functional cell analyses using the putative inhibitors that I discovered in this research, which may help elucidate the contribution of PARP14 and DDX41 to STAT6 signaling and other functions, respectively. The enzymological research described here is a key step for drug discovery. The results can be significant for discovering novel drugs against STAT6 related diseases and understanding the mechanism of STAT6 signaling.

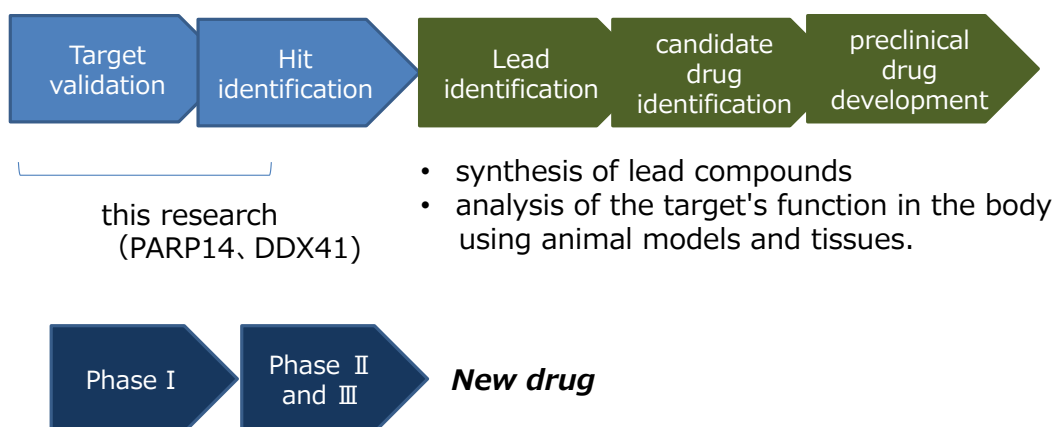


Figure 20 The process of drug discovery and development

## Acknowledgements



I am deeply grateful to Professor Tetsuo Hashimoto and Professors Kenji Miura, Kisaburo Namune, and Yuji Inagaki, University of Tsukuba, for guiding my work and valuable discussions through my doctoral program.

I also thank Mr. Shinichi Matsumoto, Takeda Pharmaceutical Company Limited, for valuable suggestions, contributions and helpful supports.

I also thank Mr. Keisuke Watanabe and Akira Tokunaga, Takeda Pharmaceutical Company Limited, for their understanding and support on my doctoral program.

Finally, I would like to appreciate my family for supporting my life in University of Tsukuba.

## References

1. Akira, S., *Innate immunity and adjuvants*. Philos Trans R Soc Lond B Biol Sci, 2011. **366**(1579): p. 2748-55.
2. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition*. Annu Rev Immunol, 2002. **20**: p. 197-216.
3. Blazquez, A.B., J. Cuesta, and C. Mayorga, *Role of dendritic cells in drug allergy*. Curr Opin Allergy Clin Immunol, 2011. **11**(4): p. 279-84.
4. Yam-Puc, J.C., et al., *Role of B-cell receptors for B-cell development and antigen-induced differentiation*. F1000Research, 2018. **7**: p. 429.
5. Pieper, K., B. Grimbacher, and H. Eibel, *B-cell biology and development*. J Allergy Clin Immunol, 2013. **131**(4): p. 959-71.
6. Burrows PD, C.M., *<B cell development and differentiation.pdf>*. Curr Opin Immunol, 1997. **2**(9): p. 239-244.
7. O'Shea, J.J. and W.E. Paul, *Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells*. Science, 2010. **327**(5969): p. 1098-102.
8. Sakaguchi, S., et al., *Regulatory T cells and immune tolerance*. Cell, 2008. **133**(5): p. 775-87.
9. Cavalla, F., et al., *Cytokine Networks Regulating Inflammation and Immune Defense in the Oral Cavity*. Current Oral Health Reports, 2014. **1**(2): p. 104-113.
10. Zhang, J.M. and J. An, *Cytokines, inflammation, and pain*. Int Anesthesiol Clin, 2007. **45**(2): p. 27-37.
11. Tamiya, T., et al., *Suppressors of cytokine signaling (SOCS) proteins and JAK/STAT pathways: regulation of T-cell inflammation by SOCS1 and SOCS3*. Arterioscler Thromb Vasc Biol, 2011. **31**(5): p. 980-5.
12. O'Shea, J.J., et al., *The JAK-STAT pathway: impact on human disease and*

- therapeutic intervention*. Annu Rev Med, 2015. **66**: p. 311-28.
13. Richard, A.J. and J.M. Stephens, *The role of JAK-STAT signaling in adipose tissue function*. Biochim Biophys Acta, 2014. **1842**(3): p. 431-9.
  14. Ingley, E. and S.P. Klinken, *Cross-regulation of JAK and Src kinases*. Growth Factors, 2006. **24**(1): p. 89-95.
  15. Takeda, K., et al., *Essential role of Stat6 in IL-4 signalling*. Nature, 1996. **380**(6575): p. 627-30.
  16. Shimoda, K., et al., *Lack of IL-4-induced Th2 response and IgE class switching in mice with disrupted Stat6 gene*. Nature, 1996. **380**(6575): p. 630-3.
  17. Hebenstreit, D., et al., *Signaling mechanisms, interaction partners, and target genes of STAT6*. Cytokine Growth Factor Rev, 2006. **17**(3): p. 173-88.
  18. Mikita, T., et al., *Mutational analysis of the STAT6 SH2 domain*. J Biol Chem, 1998. **273**(28): p. 17634-42.
  19. Mikita, T., et al., *Requirements for interleukin-4-induced gene expression and functional characterization of Stat6*. Mol Cell Biol, 1996. **16**(10): p. 5811-20.
  20. Goenka, S. and M.H. Kaplan, *Transcriptional regulation by STAT6*. Immunol Res, 2011. **50**(1): p. 87-96.
  21. Kuperman, D.A. and R.P. Schleimer, *Interleukin-4, interleukin-13, signal transducer and activator of transcription factor 6, and allergic asthma*. Curr Mol Med, 2008. **8**(5): p. 384-92.
  22. Sehra, S., et al., *IL-4 is a critical determinant in the generation of allergic inflammation initiated by a constitutively active Stat6*. J Immunol, 2008. **180**(5): p. 3551-9.
  23. Kuperman, D., et al., *Signal transducer and activator of transcription factor 6*

- (Stat6)-deficient mice are protected from antigen-induced airway hyperresponsiveness and mucus production.* J Exp Med, 1998. **187**(6): p. 939-48.
24. Goenka, S., S.H. Cho, and M. Boothby, *Collaborator of Stat6 (CoaSt6)-associated poly(ADP-ribose) polymerase activity modulates Stat6-dependent gene transcription.* J Biol Chem, 2007. **282**(26): p. 18732-9.
  25. Mehrotra, P., et al., *PARP-14 functions as a transcriptional switch for Stat6-dependent gene activation.* J Biol Chem, 2011. **286**(3): p. 1767-76.
  26. Wang, L., et al., *PARP1 in Carcinomas and PARP1 Inhibitors as Antineoplastic Drugs.* Int J Mol Sci, 2017. **18**(10).
  27. Singh, S.S., et al., *A review on PARP1 inhibitors: Pharmacophore modeling, virtual and biological screening studies to identify novel PARP1 inhibitors.* Curr Top Med Chem, 2014. **14**(17): p. 2020-30.
  28. Malyuchenko, N.V., et al., *PARP1 Inhibitors: antitumor drug design.* Acta Naturae, 2015. **7**(3): p. 27-37.
  29. Chen, H., et al., *Activation of STAT6 by STING is critical for antiviral innate immunity.* Cell, 2011. **147**(2): p. 436-46.
  30. Oliver, J.A., et al., *IL-4 acts as a potent stimulator of IFN-gamma expression in CD8+ T cells through STAT6-dependent and independent induction of Eomesodermin and Tbet.* Cytokine, 2012. **57**(1): p. 191-9.
  31. Osuala, K.O. and B.F. Sloane, *Many Roles of CCL20: Emphasis on Breast Cancer.* Postdoc J, 2014. **2**(3): p. 7-16.
  32. Yadav, A., V. Saini, and S. Arora, *MCP-1: chemoattractant with a role beyond immunity: a review.* Clin Chim Acta, 2010. **411**(21-22): p. 1570-9.
  33. Zhang, Z., et al., *The helicase DDX41 senses intracellular DNA mediated by the*

- adaptor STING in dendritic cells*. Nat Immunol, 2011. **12**(10): p. 959-65.
34. Lee, K.G., et al., *Bruton's tyrosine kinase phosphorylates DDX41 and activates its binding of dsDNA and STING to initiate type 1 interferon response*. Cell Rep, 2015. **10**(7): p. 1055-65.
  35. Gibson, B.A. and W.L. Kraus, *New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs*. Nat Rev Mol Cell Biol, 2012. **13**(7): p. 411-24.
  36. Barkauskaite, E., et al., *The recognition and removal of cellular poly(ADP-ribose) signals*. FEBS J, 2013. **280**(15): p. 3491-507.
  37. Daugherty, M.D., et al., *Rapid evolution of PARP genes suggests a broad role for ADP-ribosylation in host-virus conflicts*. PLoS Genet, 2014. **10**(5): p. e1004403.
  38. Kleine, H., et al., *Substrate-assisted catalysis by PARP10 limits its activity to mono-ADP-ribosylation*. Mol Cell, 2008. **32**(1): p. 57-69.
  39. Rosenthal, F., et al., *Macrodomain-containing proteins are new mono-ADP-ribosylhydrolases*. Nat Struct Mol Biol, 2013. **20**(4): p. 502-7.
  40. Feijs, K.L., et al., *Macrodomain-containing proteins: regulating new intracellular functions of mono(ADP-ribosyl)ation*. Nat Rev Mol Cell Biol, 2013. **14**(7): p. 443-51.
  41. Karras, G.I., et al., *The macro domain is an ADP-ribose binding module*. EMBO J, 2005. **24**(11): p. 1911-20.
  42. Mehrotra, P., et al., *Poly (ADP-ribose) polymerase 14 and its enzyme activity regulates T(H)2 differentiation and allergic airway disease*. J Allergy Clin Immunol, 2013. **131**(2): p. 521-31 e1-12.
  43. Cho, S.H., et al., *Glycolytic rate and lymphomagenesis depend on PARP14, an ADP ribosyltransferase of the B aggressive lymphoma (BAL) family*. Proc Natl Acad Sci U S A, 2011. **108**(38): p. 15972-7.

44. Wahlberg, E., et al., *Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors*. Nat Biotechnol, 2012. **30**(3): p. 283-8.
45. Andersson, C.D., et al., *Discovery of ligands for ADP-ribosyltransferases via docking-based virtual screening*. J Med Chem, 2012. **55**(17): p. 7706-18.
46. Peng, B., et al., *Small Molecule Microarray Based Discovery of PARP14 Inhibitors*. Angew Chem Int Ed Engl, 2017. **56**(1): p. 248-253.
47. Ekblad, T., et al., *Towards small molecule inhibitors of mono-ADP-ribosyltransferases*. Eur J Med Chem, 2015. **95**: p. 546-51.
48. Pingyuan Wang, J.L., Xue Jiang, Zhiqing Liu, Na Ye, Youjun Xu, Guangfu Yang, Yechun Xu, Ao Zhang, *Palladium-catalyzed N-arylation of 2-aminobenzothiazole-4-carboxylates/carboxamides: facile synthesis of PARP14 inhibitors*. Tetrahedron, 2014. **70**: p. 5666-5673.
49. Lowe, D.M., et al., *Lead discovery for human kynurenine 3-monooxygenase by high-throughput RapidFire mass spectrometry*. J Biomol Screen, 2014. **19**(4): p. 508-15.
50. Meng, J., et al., *Screening of HIV-1 Protease Using a Combination of an Ultra-High-Throughput Fluorescent-Based Assay and RapidFire Mass Spectrometry*. J Biomol Screen, 2015. **20**(5): p. 606-15.
51. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Res, 2004. **32**(5): p. 1792-7.
52. Gouy, M., S. Guindon, and O. Gascuel, *SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building*. Mol Biol Evol, 2010. **27**(2): p. 221-4.
53. Stamatakis, A., *RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses*

- with thousands of taxa and mixed models*. Bioinformatics, 2006. **22**(21): p. 2688-90.
54. Navaza, J. and P. Saludjian, [33] *AMoRe: An automated molecular replacement program package*. Methods Enzymol, 1997. **276**: p. 581-594.
  55. Winn, M.D., et al., *Overview of the CCP4 suite and current developments*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
  56. Emsley, P., et al., *Features and development of Coot*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
  57. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.
  58. Wang, Y. and D.E. Levy, *Comparative evolutionary genomics of the STAT family of transcription factors*. JAKSTAT, 2012. **1**(1): p. 23-33.
  59. Goldberg, A.L. and J.F. Dice, *Intracellular protein degradation in mammalian and bacterial cells*. Annu Rev Biochem, 1974. **43**(0): p. 835-69.
  60. Pantoliano, M.W., et al., *High-density miniaturized thermal shift assays as a general strategy for drug discovery*. J Biomol Screen, 2001. **6**(6): p. 429-40.
  61. Lomenick, B., R.W. Olsen, and J. Huang, *Identification of direct protein targets of small molecules*. ACS Chem Biol, 2011. **6**(1): p. 34-46.
  62. Schulze, J., et al., *Cell-based protein stabilization assays for the detection of interactions between small-molecule inhibitors and BRD4*. J Biomol Screen, 2015. **20**(2): p. 180-9.
  63. Auld, D.S., et al., *Examining Ligand-Based Stabilization of Proteins in Cells with MEK1 Kinase Inhibitors*. Assay Drug Dev Technol, 2015. **13**(5): p. 266-76.
  64. Martinez Molina, D., et al., *Monitoring drug target engagement in cells and tissues*



- using the cellular thermal shift assay*. Science, 2013. **341**(6141): p. 84-7.
65. Hall, M.P., et al., *Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate*. ACS Chem Biol, 2012. **7**(11): p. 1848-57.
  66. Goswami, R., et al., *STAT6-dependent regulation of Th9 development*. J Immunol, 2012. **188**(3): p. 968-75.
  67. Barbarulo, A., et al., *Poly(ADP-ribose) polymerase family member 14 (PARP14) is a novel effector of the JNK2-dependent pro-survival signal in multiple myeloma*. Oncogene, 2013. **32**(36): p. 4231-42.
  68. Cho, S.H., et al., *PARP-14, a member of the B aggressive lymphoma family, transduces survival signals in primary B cells*. Blood, 2009. **113**(11): p. 2416-25.
  69. Iqbal, M.B., et al., *PARP-14 combines with tristetraprolin in the selective posttranscriptional control of macrophage tissue factor expression*. Blood, 2014. **124**(24): p. 3646-55.
  70. Iansante, V., et al., *PARP14 promotes the Warburg effect in hepatocellular carcinoma by inhibiting JNK1-dependent PKM2 phosphorylation and activation*. Nat Commun, 2015. **6**: p. 7882.
  71. Polprasert, C., et al., *Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms*. Cancer Cell, 2015. **27**(5): p. 658-70.
  72. Kadono, M., et al., *Biological implications of somatic DDX41 p.R525H mutation in acute myeloid leukemia*. Exp Hematol, 2016. **44**(8): p. 745-754 e4.
  73. Schutz, P., et al., *Comparative structural analysis of human DEAD-box RNA helicases*. PLoS One, 2010. **5**(9).
  74. Muramatsu, H., *Genetic predisposition to pediatric myeloid malignancies*. Rinsho Ketsueki, 2016. **57**(6): p. 730-5.

75. Cardoso, S.R., et al., *Germline heterozygous DDX41 variants in a subset of familial myelodysplasia and acute myeloid leukemia*. Leukemia, 2016. **30**(10): p. 2083-2086.
76. Antony-Debre, I. and U. Steidl, *Functionally relevant RNA helicase mutations in familial and sporadic myeloid malignancies*. Cancer Cell, 2015. **27**(5): p. 609-11.
77. Gustafson, E.A. and G.M. Wessel, *DEAD-box helicases: posttranslational regulation and function*. Biochem Biophys Res Commun, 2010. **395**(1): p. 1-6.
78. Abdelhaleem, M., *RNA helicases: regulators of differentiation*. Clin Biochem, 2005. **38**(6): p. 499-503.
79. Linder, P. and E. Jankowsky, *From unwinding to clamping - the DEAD box RNA helicase family*. Nat Rev Mol Cell Biol, 2011. **12**(8): p. 505-16.
80. Shih, J.W. and Y.H. Lee, *Human DExD/H RNA helicases: emerging roles in stress survival regulation*. Clin Chim Acta, 2014. **436**: p. 45-58.
81. Jiang, Y., et al., *The emerging roles of the DDX41 protein in immunity and diseases*. Protein Cell, 2017. **8**(2): p. 83-89.
82. Parvatiyar, K., et al., *The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response*. Nat Immunol, 2012. **13**(12): p. 1155-61.
83. Zhu, X., et al., *Molecular cloning and functional characterization of porcine DEAD (Asp-Glu-Ala-Asp) box polypeptide 41 (DDX41)*. Dev Comp Immunol, 2014. **47**(2): p. 191-6.
84. Ito, M., et al., *Discovery of selective ATP-competitive eIF4A3 inhibitors*. Bioorg Med Chem, 2017. **25**(7): p. 2200-2209.
85. Sheldrick, G.M., *A short history of SHELX*. Acta Crystallogr A, 2008. **64**(Pt 1): p. 112-22.

86. Yang, J., R.A. Copeland, and Z. Lai, *Defining balanced conditions for inhibitor screening assays that target bisubstrate enzymes*. J Biomol Screen, 2009. **14**(2): p. 111-20.
87. Young, J.C., *Mechanisms of the Hsp70 chaperone system*. Biochem Cell Biol, 2010. **88**(2): p. 291-300.
88. Pause, A., N. Methot, and N. Sonenberg, *The HRIGRXXR region of the DEAD box RNA helicase eukaryotic translation initiation factor 4A is required for RNA binding and ATP hydrolysis*. Mol Cell Biol, 1993. **13**(11): p. 6789-98.
89. Zhang, L., et al., *Development of a novel helicase assay using electrochemiluminescence*. Anal Biochem, 2001. **293**(1): p. 31-7.